

Svoluji k zapůjčení své diplomové práce ke studijním účelům a prosím, aby byla vedena přesná evidence vypůjčovateli. Převzaté údaje je vypůjčovateli povinen řádně ocitovat.

**Charles University in Prague**  
**Faculty of Science**

Study programme: Biology

Branch of study: Cell and developmental biology

**Bc. Fabián Čaja**

Mutations in *MLH1* gene and MSI status as molecular characteristics of  
sporadic colorectal cancer

*Mutace v genu MLH1 a MSI status jako molekulární charakteristiky  
sporadické formy kolorektálního karcinomu*

**DIPLOMA THESIS**



*Supervisor:* MUDr. Pavel Vodička, CSc.  
*Consultant:* RNDr. Lucie Schwarzová, Ph.D.  
Prague, 2012

This diploma thesis was elaborated in the years 2010 – 2012 at the Department of the Molecular Biology of Cancer, Institute of Experimental Medicine, AS CR in cooperation with the Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University in Prague and General University Hospital in Prague and Thomayer Hospital in Prague.

*This analysis has been done within a frame of more complex CRC investigations: GACR P304/12/1585 Molecular characteristics of DNA repair in tumour tissues of CRC patients.*

### **Declaration**

I declare that this diploma thesis was carried out by myself by using the cited literature only and under the leadership of my supervisor MUDr. Pavel Vodička, CSc. and consultant RNDr. Lucie Schwarzová, Ph.D.

Prague, May 9<sup>th</sup>, 2012

## **Acknowledgement**

I would like to thank to all labmates at the Institute of Biology and Medical Genetics at First Faculty of Medicine for very pleasant atmosphere during my work on diploma thesis, for help they have dedicated to me and for advices they have given me. My special thanks belong to my supervisor MUDr. Pavel Vodička, CSc. and my consultant RNDr. Lucie Schwarzová, Ph.D. for their valuable ideas during my work and all patience with me.

I also would like to thank to my parents and my family for their support and understanding during my study at the University.

## Abstract

Colorectal carcinoma (CRC) is one of the most prevalent malignancies in the Czech Republic. In general, there are two molecular pathways leading to CRC: one is characterized by chromosomal instability, the other by the deficiency in DNA mismatch repair (MMR) genes. *MutL homologue 1 (MLH1)* gene, a member of the MMR gene-family, represents a key component of the MMR system, responsible for recognition of nucleotide mismatches occurring during DNA replication, and for the recruitment of repair proteins to correct the replication errors. According to literature, somatic mutations in MMR genes, and *MLH1* in particular, hallmark sporadic, MMR deficient, CRC cases.

We aimed at analyzing somatic events in *MLH1* gene and the determination of microsatellite instability (MSI) status in 99 DNA samples from 96 patients with sporadic CRC. Mutations were screened by high resolution melting (HRM) curve analysis. Positive cases in each run were subsequently verified by automated sequencing.

Mainly gene variants were found in *MLH1* gene: We discovered two new variants, one in exon 2 at position c. 204 C>G, p. Ile68Met (98 C/C, 1 C/G) and the other in exon 11 at position c. 973 C>T, p. Arg325Trp (98 C/C, 1 C/T). Only the latter variant c. 973 C>T was identified as somatic mutation. All other variants found in *MLH1* gene were germ-line variants. Most predominant was polymorphism c. 655 A>G, p. Ile219Val in exon 8, found in 50 DNA samples (49 A/A, 34 A/G and 16 with G/G). The other polymorphisms were detected in intron 13 at position c. 1558+14 G>A (93 G/G, 6 G/A), in intron 14 at position c. 1668-19 A>G (23 A/A, 56 A/G, 20 G/G), in exon 17 c. 1959 G>T, p. Leu653Leu (94 G/G, 5 G/T) and in exon 19 c. 2146 G>A, p. Val716Met (97 G/G, 2 G/A). We also detected mutation in exon 16 at position c. 1733 A>G, p. Glu578Gly in one patient (98 A/A, 1 A/G), which was previously described in hereditary non-polyposis colorectal cancer (HNPCC). Microsatellite instability-high (MSI-H) status was determined in 9 DNA samples (8 patients), the other 90 samples (88 patients) were considered as microsatellite stable (MSS).

We may conclude that somatic mutations in *MLH1* gene in investigated patients apparently play minor role in the development of sporadic form of CRC.

**Key words:** *MLH1* gene, colorectal cancer, mismatch repair, mutations

## Abstrakt

Kolorektální karcinom (CRC) je jedním z nejrozšířenějších maligních onemocnění v České republice. Obecně existují dvě molekulární dráhy vedoucí k CRC: první je dráha chromozomální nestability, druhá pak souvisí s poruchou reparace chybně zpárovaných bází DNA (MMR). Gen *MutH homologue 1 (MLH1)* je jedním z členů genové rodiny opravy chybného párování bází (MMR) a představuje klíčovou součást systému MMR, zodpovědného ve spolupráci s dalšími významnými proteiny za opravu poruch párování v DNA, vzniklých během replikace. Jak vyplývá z literatury, somatické mutace v MMR genech, včetně genu *MLH1*, jsou atributem sporadické formy CRC s deficientním MMR systémem.

Cílem této práce byla analýza somatických změn v *MLH1* genu a stanovení statutu mikrosatelitové nestability (MSI) v souboru 99 DNA vzorků od 96 pacientů s diagnózou sporadické formy CRC. Mutace byly detekovány pomocí HRM techniky. Pozitivní vzorky byly následně ověřovány automatickým sekvenováním.

V genu *MLH1* byly nalezeny především genové varianty: Identifikovali jsme dvě nové varianty, jednu v exonu 2 na pozici c. 204 C>G, p. Ile68Met (98 C/C, 1 C/G) a druhou v exonu 11 na pozici c. 973 C>T, p. Arg325Trp (98 C/C, 1 C/T). Pouze varianta c. 204 C>G byla stanovena jako somatická, všechny další varianty byly tzv. „germ line“ (varianty zárodečné linie). Nejpočetněji zastoupený byl polymorfismus c. 655 A>G, p. Ile219Val v exonu 8, jež byl stanoven u 50 vzorků (49 A/A, 34 A/G a 16 G/G). Další polymorfizmy byly prokázány v intronu 13 na pozici c. 1558+14 G>A (93 G/G, 6 G/A), v intronu 14 na pozici c. 1668-19 A>G (23 A/A, 56 A/G, 20 G/G), v exonu 17 c. 1959 G>T, p. Leu653Leu (94 G/G, 5 G/T) a v exonu 19 c. 2146 G>A, p. Val716Met (97 G/G, 2 G/A). Rovněž byla identifikována mutace v exonu 16 na pozici c. 1733 A>G, p. Glu578Gly (98 A/A, 1 A/G), dříve popsaná u hereditárního nepolypózního kolorektálního karcinomu (HNPCC). MSI-H status byl stanoven u 9 vzorků (8 pacientů), ostatních 90 vzorků od 88 pacientů lze považovat za mikrosatelitně stabilní (MSS).

Lze konstatovat, že somatické mutace v *MLH1* genu v daném souboru pacientů hrají pravděpodobně minoritní roli ve vývoji sporadické formy CRC.

**Klíčová slova:** MLH1 gen, kolorektální karcinom, oprava chybného párování, mutace

# Content

## List of abbreviation

<b>1. Introduction .....</b>	<b>9</b>
<b>2. State of art .....</b>	<b>11</b>
2.1. Colorectal cancer	
2.1.1. Epidemiology of colorectal cancer (CRC).....	11
2.1.2. Main genetic characteristics of CRC .....	13
2.2. Mismatch repair (MMR) pathway in colorectal carcinogenesis	
2.2.1. MMR and sporadic colorectal cancer .....	21
2.2.2. Microsatellite instability (MSI) status and CRC .....	25
2.2.3. MSI and CIMP markers of early progression, clinical outcomes and the risk of sporadic CRC .....	28
<b>3. Aims of diploma thesis .....</b>	<b>31</b>
<b>4. Materials and methods</b>	
4.1. Study population .....	32
4.2. Ethical approval .....	32
4.3. DNA processing .....	33
4.4. Detection of MSI status .....	33
4.5. Mutational analysis of <i>MLH1</i> gene – HRM, PCR .....	34
4.6. Mutational analysis of <i>MLH1</i> gene – automated sequencing .....	39
4.7. Data analysis .....	40
<b>5. Results</b>	
5.1. Clinical characteristics of CRC patients .....	42
5.2. Mutations in <i>MLH1</i> gene .....	44
<b>6. Discussion .....</b>	<b>59</b>
<b>7. Conclusions .....</b>	<b>65</b>
<b>8. References .....</b>	<b>66</b>

## **List of abbreviation**

5-FU – 5-fluorouracil

AFAP – attenuated familial adenomatous polyposis

APC – adenomatous polyposis coli

BAX – Bcl-2 associated X protein

BCL- B-cell lymphoma

BMPR – bone morphogenic protein receptor

BRAF – rapidly accelerated fibrosarcoma, protein type B

CACNA – calcium channel gene A

CDKN – cyclin-dependent kinase

CIMP – CpG methylator phenotype

CIN – chromosomal instability

CRC – colorectal cancer

CS – Cowden syndrome

DCC – deleted in colorectal cancer

DMSO – dimethyl sulphoxide

DNMT – DNA methyl transferase

EXO – exonuclease

FAP – familial adenomatous polyposis

FCC – familial colorectal cancer

GTP – guanosin triphosphate

HMPS – hereditary mixed polyposis syndrome

HNPCC – hereditary non-polyposis colorectal cancer

HRM – high resolution melting

HPPS – hyperplastic polyposis syndrome

IDL – insertion/deletion loop

IGF – insuline growth factor

IGFR – insuline growth factor receptor

JPS – juvenile polyposis syndrome

KRAS – Kirsten rat sarcoma



LOH – loss of heterozygosity  
LS – Lynch syndrome  
MAPK – mitogen activated protein kinase  
MINT – methylated in tumours  
MLH1 – MutL homologue 1  
MMP – matrix metalloproteinase  
MMR – mismatch repair  
MSH – MutS homologue  
MSI – microsatellite unstable  
MSS – microsatellite stable  
MUTYH – Mut Y homologue  
NEUROG – neurogenin  
PCR – polymerase chain reaction  
PJS - Peutz-Jeghers syndrome  
PMS – postmeiotic segregation increased  
PTEN – phosphatase and tensin homologue  
RAD – Ras associated with diabetes  
RAS – Rat sarcoma protein  
RUNX - Runt-related transcription factor  
SMAD – homologue of SMA and mothers against decapentaplegic (MAD)  
SOCS - suppressor of cytokine signaling  
STK – serine-threonine kinase  
TACSTD - Tumor-associated calcium signal transducer  
TGF – transforming growth factor  
VEGF – vascular endothelial growth factor  
WHO – World Health Organisation  
Wnt – homologue of Wingless (Wg) and Int

## 1. Introduction

One of the main functions of living cells is to maintain genomic stability and copy its genetic information precisely into daughter cells. Accumulation of genetic alterations in the cell, such as mutations, loss of heterozygosity (LOH) and chromosomal rearrangements, may give one cell a selective advantage, resulting in more rapid growth and division and its better survival. Nowadays it is assumed that at least 6 to 12 mutations are necessary for malignant transformation of the cell (Wood et al., 2007). There are several mechanisms that control genomic stability in every cell. Regulation of the cell cycle and repair of DNA represent the most prominent trait.

Colorectal cancer (CRC) is one of the most frequent cancers in the developed countries; it is for instance the second most common form of cancer in Europe (Ferlay et al., 2007). It poses a serious public health problem in the Czech Republic too; this country had the second highest CRC incidence and mortality among 38 European countries in 2007 (Ferlay et al., 2007). Both genetic and environmental factors may underlie frequent occurrence of sporadic CRC.

We distinguish both hereditary and sporadic forms of cancer. Hereditary forms of CRCs, accounting for about 5 % of all CRCs, develop at a young age due to germ line mutations in crucial genes. About 80 % of CRCs are sporadic forms which develop due to somatic mutations in important regulatory genes or their deregulation by epigenetic mechanisms, mainly by promoter hypermethylation (Boland and Goel, 2010). Dietary, life-style and environmental factors play important role in the development of sporadic CRC. Sporadic form of CRC may arbitrary be divided into two pathways: chromosomal instability (CIN) and microsatellite instability (MSI) pathways. The former, accounting for 80-85 % of all sporadic CRCs (Lengauer et al, 1998) proceeds through mis-segregation of chromosomes that results in aneuploidy and through unbalanced structural rearrangements leading to the loss and/or gain of chromosomal regions (Rajagopalan et al., 2003). The latter one accounts for about 15-20 % of sporadic cases of CRCs (Boland and Goel, 2010). Microsatellites are short repetitive DNA sequences in all regions of the genome. Insertions/deletions in replicated DNA, including

microsatellites, are repaired by MMR proteins such as MutH homologue 1 (MLH1), MutS homologue 2 (MSH2), Postmeiotic segregation increased 1 (PMS1), PMS2, MSH6 and MLH3. Defective MMR proteins, which cannot repair alterations in microsatellites caused by erroneous activity of DNA polymerase, are manifested as microsatellite instability (MSI) phenotype.

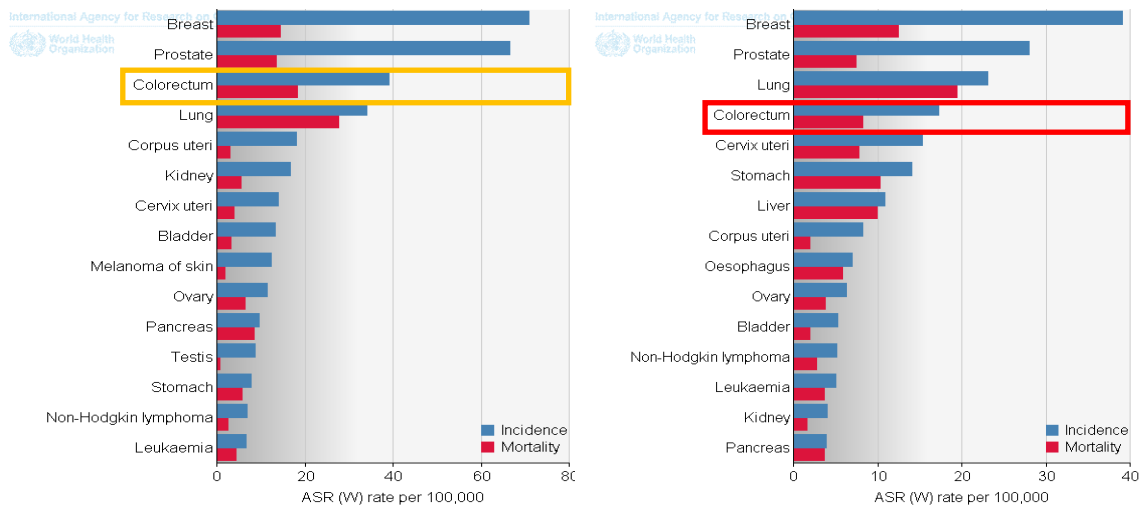
The aim of the present study was to investigate the somatic events in high-risk MMR gene *MLH1* in patients with the sporadic form of CRC. Present study is one of the first genetic analyzes of *MLH1* gene on the Czech population. As mentioned above, Czech Republic has one of the highest incidencies of CRC in Europe, thus understanding the principles of MMR process may ultimately contribute the effectiveness of anticancer therapy.

## 2. State of art

### 2.1. Colorectal cancer

#### 2.1.1. Epidemiology of colorectal cancer

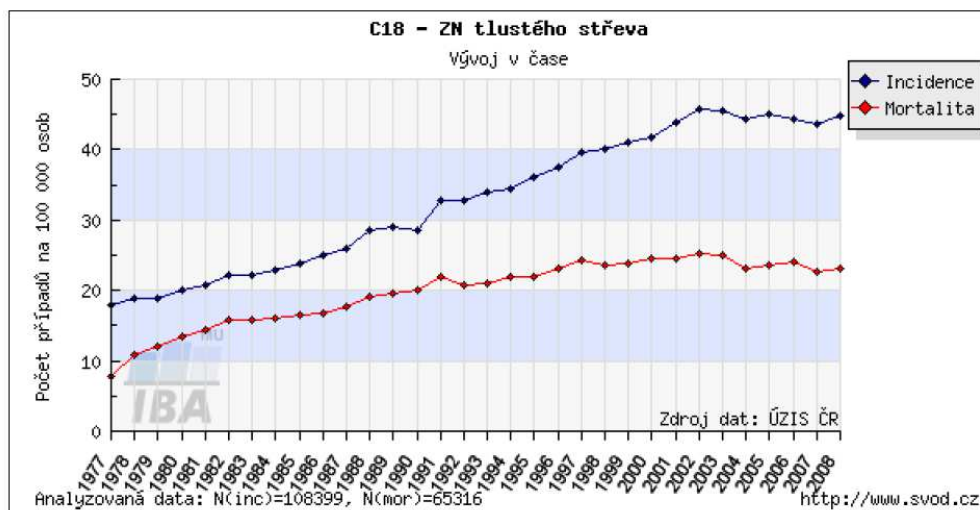
Globally more than 1 million people get CRC annually, resulting in about 0.5 million deaths (Cunningham et al., 2010). CRC is the fourth most common type of cancer death after breast, prostate and lung cancer (**Figure 1**) (Jemal et al., 2011; <http://www.who.int/en>). Around 60% of cases are diagnosed in the developed world (<http://www.globocan.com>). CRC accounts for 12.9 % of all malignant disorders in Europe (Ferlay et al., 2007). For illustration, there were 412,000 new CRC cases diagnosed and 207,000 deaths caused by CRC in Europe during 2006. Despite rising incidence of CRC in developed countries, the mortality of colorectal cancer is decreasing mainly due to early screening.



**Figure 1: Incidence and mortality of selected types of cancer** in the Czech Republic in the left graph and worldwide in the right graph. CRC is the third most common type of cancer in the Czech Republic and the fourth one worldwide. Taken from <http://www.globocan.com>.

CRC poses a serious health problem in the Czech Republic too (**Figure 1, Figure 2 - p. 12**). According to National Cancer Registry (<http://www.svod.cz>), it is the second most prevalent type of cancer in men and the third most prevalent type of cancer in women with 52.5 and 37.6 new cases, respectively (in 2008 per 100,000

subjects, <http://www.svod.cz>). Every year 7,900 – 8,100 new cases are diagnosed and 4,300 – 4,500 patients die in this country (Dušek et al., 2005). The mortality rate is one of the highest in the European region, caused by late diagnosis which is connected with worse prediction. An average age of people with the CRC diagnosis in the Czech Republic ranges from 60 to 75, but 23 % of patients are younger than 60 (<http://www.kolorektum.cz>). A high fat, alcohol or red meat intake are risk factors for CRC as well as it is obesity, smoking and a lack of physical exercise (Watson and Collins, 2011). The high incidence of CRC in the Czech Republic may also be a consequence of the accumulation of more genetic variants in the gene pool predisposing to CRC, for example single nucleotide polymorphisms (SNPs). The role of genetic polymorphisms in *ATP-binding cassette receptor G2 (ABCG2)* gene was investigated in the Czech population. Heterozygous carriers of the minor alleles of SNPs rs2622621 and rs1481012 had a decreased risk of CRC (Campa et al., 2008). Similar study investigated the association between genetic polymorphisms in genes coding for ghrelin (GHRL) and its receptor (GHSR) and CRC susceptibility (Campa et al., 2010). Two SNPs, namely SNPs rs27647 and rs35683, were associated with lower risk of CRC. Due to the worldwide severity of the problem, colorectal carcinogenesis is one of the most extensively studied types of cancer.



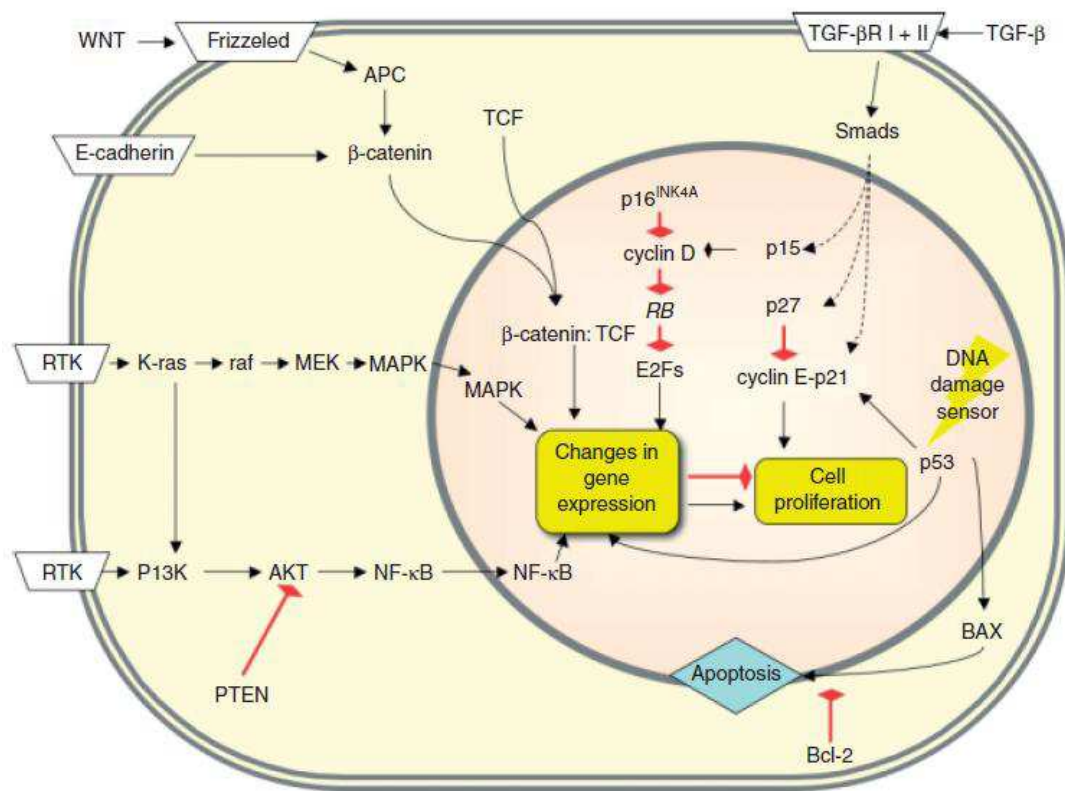
**Figure 2: Incidence and mortality of CRC in the Czech Republic from 1977 to 2008.**  
Blue line represents incidence and red one mortality. Adapted from <http://www.svod.cz>.

CRC is developing through a series of histopathological and clinical stages ranging from small lesions in the colon crypts through benign polyps to malignant carcinoma. Tumour-node-metastasis (TNM) is used for defining these stages and it is based on depth of tumour invasion in the colon mucosa, presence of cancer cells in the regional lymph nodes and in the distant sites (Wolpin et al. 2008). Another grading system is Dukes classification, based on the percentage of the differentiated cells (Labianca et al., 2010). The cure strategy and prognosis is based upon these classifications.

### **2.1.2. Main genetic characteristics of CRC**

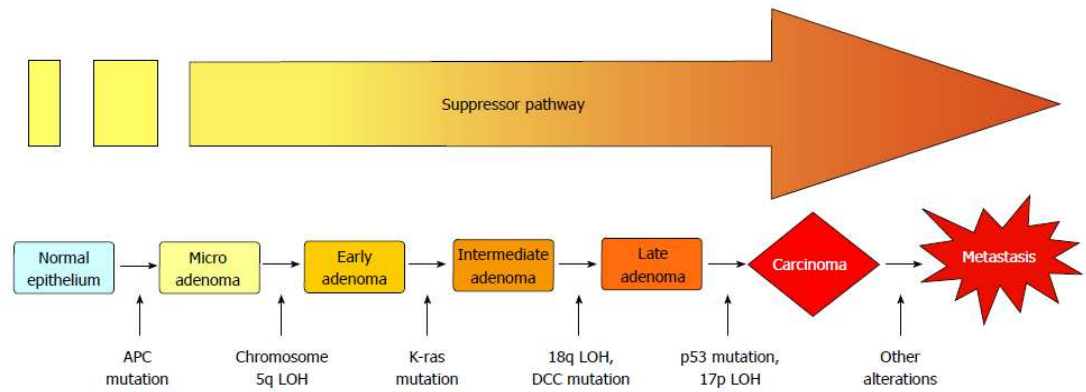
Generally there are two groups of CRCs resulting from gene-environment interactions: approximately 80% of colorectal tumours have a sporadic origin, remaining 20 % have a family history (Jaspersion et al., 2010). However, only 5-6 % of CRC is due to inherited mutations in major CRC genes (Jaspersion et al., 2010).

Current opinion is that there are two major pathways in sporadic colorectal carcinogenesis. Chromosomal instability (CIN) pathway, also called “suppressor” pathway, is characterized by allelic losses on chromosomes 5q in *Adenomatous polyposis coli* (APC) gene, 17p (*p53* gene) and 18q in *Deleted in colorectal cancer* and *homologue of SMA and mothers against decapentaplegic* (MAD) 4 gene (*DCC* and *SMAD4*) (Laurent-Puig et al, 2010). Microsatellite instability (MSI) pathway, also called “mutator” pathway, is characterized by genetic and epigenetic inactivation of DNA mismatch repair (MMR) genes and by accumulation of mutations in microsatellite sequences. These mutations are caused by erroneous activity of DNA polymerase during replication of DNA (Imai and Yamamoto, 2008). Complexity and tentative cross-talk between various pathways is illustrated in **Figure 3, p. 14.**



**Figure 3: Simplified overview of general pathways in colorectal cancer.** Black arrows suggest a positive interaction; red bars suggest negative control. WNT denotes *Wingless* pathway. AKT, a serine/threonine kinase; APC, adenomatous polyposis coli; BAX, Bcl-2-associated X protein; E2F, E2F transcription factors; MAPK, mitogen-activated protein kinase; MEK, MAPK kinase; NF, nuclear factor; PI3K, phosphatidylinositol 3-kinase; PTEN, phosphatase and tensin homologue; *RB*, *retinoblastoma* gene; RTK, receptor tyrosine kinase; SMAD, mothers against decapentaplegic homologue (*Drosophila*); TCF, T cell factor; TGF-β(R), transforming growth factor β (receptor). Adapted from Søreide et al., 2006.

The basic model of colorectal carcinogenesis (**Figure 4, p. 15**), proposed by Fearon and Vogelstein in 1990, is still valid despite accumulation of new knowledge. It is thought that in majority of cases carcinomas arise from pre-existing adenomas. Every step in this process involves a specific and well-defined genetic alterations (Fearon et Vogelstein, 1990). This pathway is characterized by alterations in tumour suppressor genes, such as *APC*, *p53* and *DCC* and in oncogenes, such as *Kirsten rat sarcoma* (*KRAS*). All these genes are discussed in next paragraphs. The “CIN” tumours are also characterized by a high frequency of allelic imbalance, chromosomal amplifications and translocations (Pawlik et al., 2004).



**Figure 4: Molecular alterations in the CIN pathway.** The basic scheme of CIN pathway was proposed by Fearon and Vogelstein. The main molecular mechanisms are still valid despite new discoveries. The first mutational events in APC gene lead to adenoma transition followed by other molecular mechanisms to adenoma-carcinoma transition caused mainly by p53 deficiency. Typical feature of CIN pathway is huge chromosomal imbalance caused by chromosomal amplifications and translocations. Adapted from Morán et al., 2010.

**APC:** The *adenomatous polyposis coli* (APC) gene, located at 5q21 chromosome, contains 15 exons and it is mutated in 60 % of colon and 82 % of rectal cancers, respectively (Jass, 2002). The APC protein has several functions in the cellular physiology, but its best-known role is in the Wnt pathway. APC is a part of multiprotein complex that joins  $\beta$ -catenin and causes its phosphorylation, ubiquitination and subsequent degradation in the proteasome. The mutations in APC gene interfere with its binding to  $\beta$ -catenin and result in the Wnt pathway becoming constitutively active. The  $\beta$ -catenin can translocate into the cell nucleus where it regulates the transcription of crucial regulatory genes related to tumour progression, such as *Cyclin D*, *c-Myc*, *Matrix metalloproteinase 7* (MMP-7) and *MMP-26* (Barker, 2008). In 1971, Alfred Knudson formulated the hypothesis that cancer is the result of accumulated mutations to a cellular DNA. However, most loss-of-function mutations that occur in tumour suppressor genes are recessive in nature. Thus, both of the cellular tumour suppressor genes must be mutated to develop a cancerous cell. This idea is known as the "two-hit" hypothesis (Knudson, 1971). Vasovcak et al. reported that frequency of APC defects rose to 75.7 % if LOH at the APC locus was included. Significantly increased LOH in tumours with just 1 somatic mutation supports the two-hit model of tumorigenesis (Vasovcak et al., 2011). APC plays also other roles in cellular physiology independent of Wnt (homologue of Wingless (Wg)



and Int signalling) pathway. It participates in the regulation of the cell cycle. APC has been found at kinetochores where it controls correct chromosomal alignment, and at centrosomes, where it influences their duplication (Rusan et al., 2008).

**KRAS:** *KRAS* is a proto-oncogene, located at 12p12.1, and its 21-kDa protein belongs to the large family of small GTPases, which are activated in response to various extracellular stimuli. *KRAS* mutations, especially in codons 12 and 13 occur primarily in pancreatic and colon cancers (Kranenburg 2005) and may be detected during the very early stages of CRC development (32-42 % of sporadic CRC cases) (Leslie et al., 2002). *KRAS* is involved in many different processes, such as mitogen-activated protein kinase (MAPK) signalling pathway and other signalling pathways regulating the expression of Cyclin D1, DNA methyltransferase (DNMT) and vascular endothelial growth factor (VEGF) (Estève et al., 2005; Oren and Rotter, 2010; Rocha et al., 2003). It also regulates epithelial cell polarity. During the development of CRC, epithelial cells lose their polarity and it has been demonstrated that an acquired mutation in *KRAS* gene reduces adherens junction-mediated cell-cell contacts (Smakman et al., 2005). Mutant *KRAS* promotes hyperplastic growth in the colonic epithelium and suppresses differentiation in *APC*-mutant colon cancers (Haigis et al., 2008).

**DCC:** *Deleted in colorectal cancer (DCC)* gene is a tumour suppressor gene, located at 18q21.1, and is mutated in approximately 70 % of CRCs. DCC protein has more functions in the cell physiology. It is a transmembrane receptor of the Ig superfamily for netrins, factors involved in the developing nervous system (Livesey and Hunt, 1997). It has a role in intracellular signalling too. Wild-type DCC activates caspase-3, induces apoptosis and rapid G2/M cell cycle arrest (Chen et al., 1999). Shekarabi and Kennedy discovered that DCC activate Rac-1 in rodent neuroblastoma cells (NG108-15) and human embryonic kidney cells (HEK293T). Rac-1 is involved in actin organisation and cell motility, thus its activation may play important role in epithelial differentiation (Shekarabi and Kennedy, 2002).

**p53:** The most of human cancers have deficient p53 protein, which is encoded by tumour suppressor gene *TP53* located on 17p13.1. Many studies have shown that the frequency of p53 abnormalities increases with the progression of the

lesion and it is estimated that 50-75 % of sporadic CRCs display mutations in *TP53* (Leslie et al., 2002). Mutations or loss of *TP53* usually occurs at the time of the transition from adenoma to carcinoma. *TP53* variants, mainly in the form of SNPs, may impact susceptibility to CRC. The carriers of the variant A-allele for the rs17884306 polymorphism in exon 11 (-363G>A) were found to be at a decreased risk of rectal cancer (Polakova et al., 2009). P53 protein plays important role in the maintenance of genomic integrity. In the interaction with other pathways, p53 responds to genotoxic stress, DNA damage and nutrition stress. P53 induces G1 cell cycle arrest to facilitate DNA repair during replication. If DNA damage is too extensive to be repaired, it can induce apoptosis (Takayama et al., 2006). Mutations in *TP53* gene are more common in distal colon and rectal cancer than in proximal colon cancer. Patients bearing mutations in *TP53* gene have worse prognosis of cure than patients missing mutations in *TP53* (Russo et al., 2005). Mutations and polymorphisms in *TP53* gene with relevance for CRC have recently been overviewed (Naccarati et al., 2012).

**TGFβ/SMAD:** The growth inhibition of many cells, such as epithelial cells, is related to signal transduction by transforming growth factor-β (TGF-β) and its two receptors, TGF-β RI and TGF-β RII. Inactivation of these receptors leads to silencing of TGF-β signalling, thus several carcinoma cells show mutations and inactivation of TGF-β RI or TGF-β RII (MacKay et al., 1995). TGF-β plays other roles in cell as well, such as regulation of their migration, adhesion, differentiation and apoptosis. Therefore TGF-β signalling is important tumour suppressor pathway. TGF-β RII binds TGF-β ligand leading to its interaction with TGF-β RI which then phosphorylates SMAD 2 and SMAD 3. It leads to formation of SMAD 2(3)/4 dimer which translocates to cell nucleus and regulate expression of CDK inhibitors p15 and p21 (Masagué et al., 2005). P21 protein expression is also activated by p53 if the chromosomal DNA of a cell suffers some damage during the G1 phase of the cell cycle (Vousden, 2002). SMAD 7 acts as an intracellular antagonist of TGF-beta signalling by binding stably to the receptor complex and blocking activation of downstream signalling events (ten Dijke and Hill, 2004). Pittman et al. identified novel SNPs in *SMAD7* gene leading to CRC predisposition through differential

SMAD 7 expression (Pittman et al., 2009). Mutations in *SMAD 4* and *SMAD 2* have been found in 16-25 % and in 6 % of CRC, respectively (Takagi et al., 1996). *TGF- $\beta$  RII* gene was reported to be highly associated with MSI-H carcinoma (Markowitz et al., 1995). Mutations in TGF- $\beta$  receptor genes occur at the time of transition from adenoma to carcinoma, like in *TP53* gene. It has been shown that components of Wnt signalling pathway interact with the components of TGF-  $\beta$  pathway mainly through proteins SMAD and lymphoid enhancer-binding factor 1 (LEF-1) (Fuentealba et al., 2007). It seems that cooperation of signalling pathways play important role in carcinogenesis (**Figure 3, p. 14**).

Hereditary CRCs develop due to inherited mutations in high penetrant genes. These mutated genes predispose individuals to the development of CRC. There are many CRC syndromes; major ones are adenomatous polyposis syndromes such as familial adenomatous polyposis (FAP), attenuated familial adenomatous polyposis (AFAP) and MUTYH-associated polyposis (MAP). Lynch syndrome (LS) also called hereditary non-polyposis colorectal cancer (HNPCC) is a non-polypoposis CRC syndrome. There were also rare hamartomous and hyperplastic polyposis CRC syndromes identified, such as Peutz-Jeghers syndrome (PJS), juvenile polyposis syndrome (JPS), Cowden syndrome and others. All these syndromes will be described in brief in the next paragraphs.

**Familial adenomatous polyposis (FAP), Attenuated FAP (AFAP):** FAP is an autosomal dominant disorder. The “second hit” is necessary for development of FAP. It is characterized by the development of thousands of adenomas in the colon or rectal mucosa in the first decade of life. Malignant transition is predicted in 100 % of adenomatous polyps. FAP accounts for less than 1 % of all CRC cases with the prevalence 1/11 300-37 600 in the European Union (Half et al., 2009). FAP may accompany other extra-intestinal pathologies such as osteomas, dental abnormalities, congenital hypertrophy of retinal pigment epithelium and extracolonic tumours (Half et al., 2009). Less aggressive variant of FAP is called attenuated FAP (AFAP), characterized by less number of adenomatous polyps and later age of adenoma

appearance (Jasperson et al., 2010). The genetic background of FAP syndrome is well known. FAP and AFAP result from germ-line mutations of adenomatous polyposis coli (*APC*) tumour suppressor gene (Benchabane and Ahmed, 2009).

**MUTYH-associated polyposis (MAP):** MAP is an autosomal recessive disorder which has very similar gastroenterological features to FAP syndrome. The disease is caused by mutations in both alleles in *MutY homologue (MUTYH)* gene located on chromosome 1p.34.1. It encodes a protein involved in DNA base excision repair (BER) pathway. Compared with the wild-type protein, the mutant proteins showed a roughly 98% (Tyr82Cys) and 86% (Gly253Asp) reduction of MUTYH activity (Al-Tassan et al., 2002).

**Lynch syndrome (LS):** An autosomal dominant disorder caused by mutations in DNA mismatch repair (MMR) genes, the most common hereditary CRC syndrome. It accounts for ~3 % of all CRCs (Lynch et al., 2009). More information about Lynch syndrome is presented further in the text, if relevant.

**Peutz-Jeghers syndrome (PJS):** It is an autosomal dominant disorder typical by hamartomatous polyps in the colon mucosa and by pigmented lesions on the lips, buccal mucosa and perioral region (Kopacova et al., 2009). Prevalence of PJS is estimated from 1/8300 to 1/ 280 000 individuals and it is often accompanied by other malignancies such as gastrointestinal, pancreatic, breast, lung, uterine, ovarian and testicular (Beggs et al., 2010). The majority of patients have a mutation in tumour suppressor gene *STK11* located at 19p13.3 chromosome, encoding a serine threonine kinase STK11 (Jenne et al., 1998).

**Juvenile polyposis syndrome (JPS):** Another hereditary CRC syndrome characterized by hamartomatous polyps in the gastrointestinal tract. It is estimated that 15-20 % of JPS patients carry autosomal dominant mutations in the *SMAD4* gene located on chromosome 18q21.1, that encodes the SMAD4 protein participating in the TGF $\beta$  signalling pathway (Howe et al., 1998). Approximately 25-40 % of JPS patients carry autosomal dominant mutations in the gene coding bone morphogenetic protein receptor 1A (*BMPR1A*) on chromosome 10q22 (Howe et al., 2001).

**Hereditary mixed polyposis syndrome (HMPS):** An autosomal dominant disorder with mixed adenomatous, hyperplastic and juvenile histology. Its phenotype is very similar to JPS. Mutations in *BMPRIA* gene contribute to development of HMPS syndrome (Cheah et al., 2009).

**Cowden syndrome (CS):** Another rare autosomal dominant disorder typical by hamartomous polyposis. Approximately 27-43 % of CS patients have hamartomous polyps in the gastrointestinal tract. The disease is caused by mutations in tumour suppressor *phosphatase and tensin homologue* gene (*PTEN*) (Liaw et al., 1997).

**Hyperplastic polyposis syndrome (HPPS):** HPPS is a rare disease, characterized by the presence of large hyperplastic polyps in the colon mucosa (Rubio et al., 2006). According to World Health Organisation (WHO), HPPS is characterized by these conditions: at least five hyperplastic polyps located in proximal to sigmoid colon, more than 30 polyps in the whole colon and at least one-degree relative with HPPS diagnosis (<http://www.who.int/en>). Little is known about the aetiology and genetics of HPPS.

Hereditary forms of CRC account for approximately 5-6 % of all CRC cases. However, it is estimated that familial clusters of CRC account for almost 20% of CRCs in developed countries and the rare hereditary CRC syndromes described above contribute only to a fraction of them (Migliore et al., 2011). The term familial colorectal cancer (FCC) is used for all hereditary CRCs which do not meet the clinical criteria for the diagnosis of known hereditary CRC syndromes. According to population studies and genome-wide association studies, low-penetrance loci and several polymorphisms were identified to increase the risk for FCC development (Jaspersion et al., 2010).

Table 1 summarizes the genes in which mutations predominantly contribute to the development of CRC (**Table 1**). Information about MMR genes are provided in the next subchapter.

<b>Gene</b>	<b>Chromosome</b>	<b>Function</b>	<b>Role</b>
APC	5q21-22	Tumour suppressor	Mutated in FAP patients
TGFβRII	3p22	Cell signalling	Associated with MSI
MSH2	2p16	DNA mismatch repair (MMR)	HNPCC
MLH1	3p21	MMR	HNPCC, epigenetically silenced in sporadic CRC
MSH6	2p16	MMR	HNPCC
PMS2	7p22	MMR	HNPCC; rare
PMS1	2q31-33	MMR	HNPCC ?
MLH3	14q24.3	MMR	HNPCC ?
KRAS	12p12.1	Oncogene	Early event CRC
p53	17p13	Tumour suppressor	Late event CRC
BAX	19q13.3-4	Apoptosis	Associated with MSI
MGMT	10q26	DNA-repair	CIMP phenotype, MSI
PTEN	10q23	Tumour suppressor	Associated with MSI

**Table 1: Selected genes involved in CRC.** APC, adenomatous polyposis coli; TGFβRII, transforming growth factor β receptor II; MSH, MutS homologue; MLH, MutL homologue; PMS, postmeiotic segregation increased; COX, cyclo-oxygenase; Bcl, B-cell chronic lymphocytic leukaemia/lymphoma; BAX, Bcl-2 associated X protein; MGMT, O-6-methylguanine DNA methyltransferase; PTEN, phosphatase and tensin homologue; CIMP, CpG island methylator phenotype. The table was adapted from Søreide et al., 2006.

## 2.2. Mismatch repair (MMR) pathway in colorectal carcinogenesis

### 2.2.1. MMR and colorectal cancer

The S-phase of the cell cycle is characterized by DNA replication owing to duplication of whole cellular genome. Replicative polymerases may incorporate wrong nucleotides leading to base-base mismatches or they can slip on nucleotide sequences creating insertion-deletion loops (IDL). The result of this erroneous activity of DNA polymerase is the production of single nucleotide point mutations or a frame shift mutations leading to synthesis of truncated protein. These biosynthetic errors in newly synthesized DNA are removed by mismatch repair (MMR) system. Several studies have revealed that loss of MMR leads to development of mutator

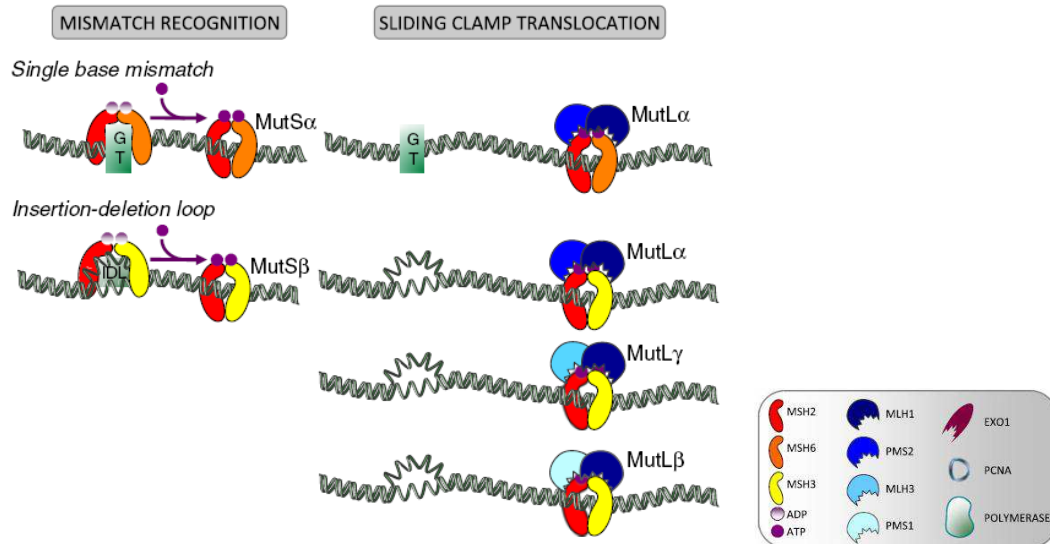
phenotype, which causes a predisposition to cancers of the colon, endometrium and other organs (Peltomaki, 2003; Marra and Jiricny, 2005).

The principles of MMR system were first understood on *E.coli* bacteria (Nevers and Spatz, 1975). It involves three specialized proteins: MutS, MutL and MutH (Lahue et al., 1989). MutS is an ATPase acting as a homodimer which recognizes base-base mismatches and small IDLs (Allen et al., 1997; Modrich and Lahue, 1996). Its loading onto the DNA leads to conformational change allowing its interaction with MutL - an ATP-dependent homodimer, which is important for removal of DNA fragment containing the mismatch (Ban et al., 1999). The newly synthesized DNA strands lack the methylation at GATC sequences, which are recognized by MutH endonuclease (Welsh et al., 1987).

Human MMR system consists of five MutS homologues (MSH2-MSH6) and four MutL homologues (MLH1, MLH3, PMS1, PMS2) (**Figure 5, p. 23**). MSH2 forms a heterodimer with MSH6 or MSH3 giving rise to the assembly of MutS $\alpha$  and MutS $\beta$  complexes (**Figure 5**), respectively (Genschel et al., 1998). It was revealed that MutS $\beta$  complexes preferentially repair heteroduplexes with larger IDLs than MutS $\alpha$  complexes (Marsischky et al., 1996). In humans, MLH1 is the major MutL homologue which assembles to MutL $\alpha$ , MutL $\beta$  and MutL $\gamma$  complexes (**Figure 5**) with PMS2, PMS1 and MLH3 proteins, respectively (Kondo et al., 2001). The basic principles of human MMR system are depicted on **Figure 6 (p. 24)**. First, MutS $\alpha$  or MutS $\beta$  complexes bind to DNA as sliding clamps after MSH2 phosphorylation. Then heterodimer MutL $\alpha$  binds to MutS $\alpha$  forming MutS $\alpha$ /MutL $\alpha$  complex. MutS $\alpha$ /MutL $\alpha$  then moves along the DNA until it meets the DNA polymerase, proliferating cell nuclear antigen (PCNA) and exonuclease 1 (EXO1). The exonuclease removes several hundred bases from new DNA strand, which allow correct re-synthesis of DNA by the replication polymerase.

MutL $\alpha$  is participating in production of nicks into DNA leading to degradation of mismatch-containing strand by EXO1, thus MutL $\alpha$  is the major component in MMR (Kadyrov et al., 2006). It seems that MutL $\gamma$  recognizes G/T mismatches and small IDLs, i.e. the same substrates as MutS $\alpha$ , whereas larger IDLs are recognized by MutL $\beta$ . MLH1 is necessary to prevent degradation of PMS2, thus

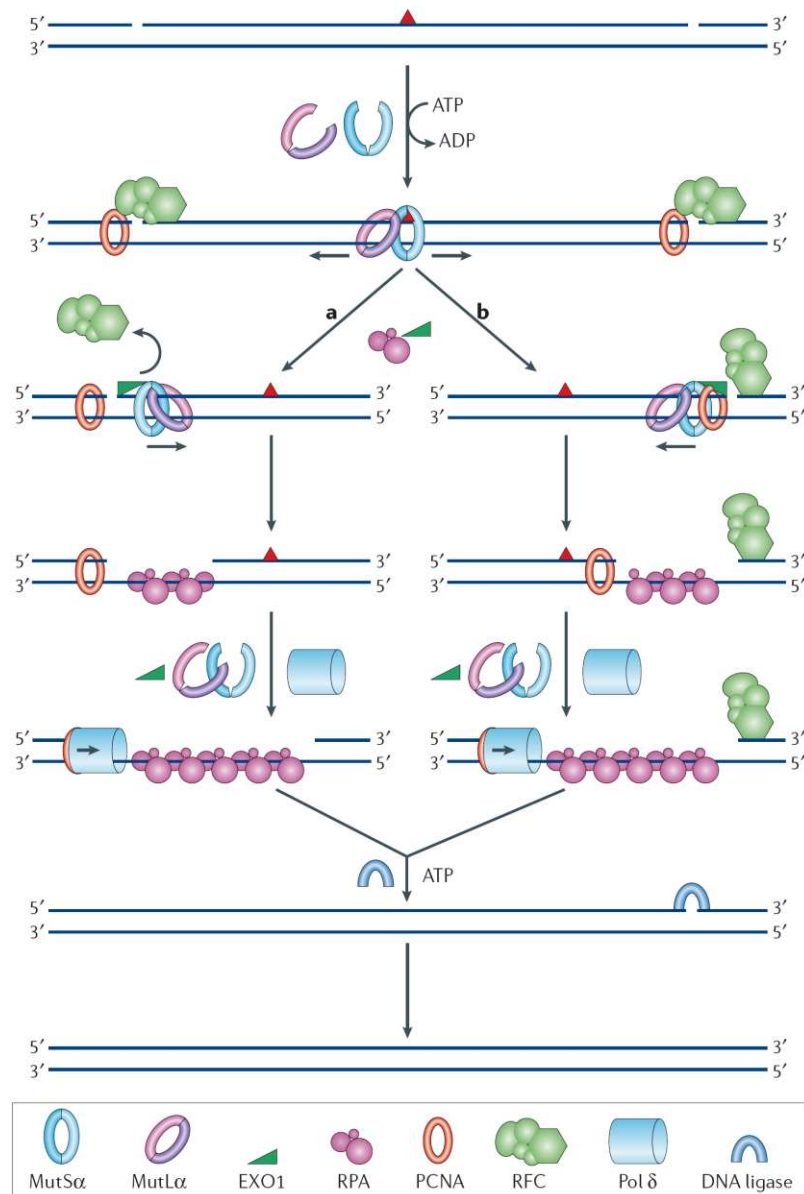
it seems that MutL $\gamma$  plays a redundant role in the absence of PMS2 which is a part of MutL $\alpha$  complex (Cannavo et al., 2005).



**Figure 5: Human MMR proteins forming MutS and MutL complexes.** Single nucleotide mismatches, as well as 1 nucleotide loops typically occurring at microsatellite mononucleotide repeated sequences, are recognized by the MutS $\alpha$  heterodimer formed by MSH2 and MSH6. IDLs of two or more nucleotides are preferentially recognized by the MutS $\beta$  complex, formed by MSH2 and MSH3. MutL $\alpha$  complex is a heterodimer formed by MLH1 and PMS2. MLH1 may also form heterodimers with MLH3 and PMS1 to form respectively the MutL $\gamma$  and MutL $\beta$  complexes. Adapted from Zaanani et al., 2011.

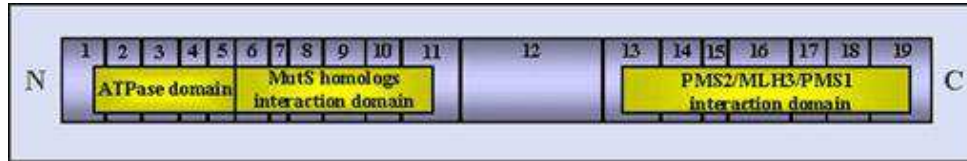
MMR complexes are also involved in DNA damage signalling, regulation of the cell cycle and p53-dependent apoptotic signalling (O'Brien and Brown, 2006; Stojic et al., 2004).





**Figure 6: The structure of human MMR system.** The MMR system consists of either MutS $\alpha$  or MutS $\beta$ , MutL, RPA, EXO1, PCNA, RFC, DNA polymerase  $\delta$  and DNA ligase I. The mismatch (red triangle)-bound MutS recruits MutL. This complex undergoes an ATP-dependent conformational switch, which releases the sliding clamp from the mismatch site. **a** | Clamps that diffuse upstream encounter RFC that is bound at the 5' terminus of the strand break, and will displace it and load EXO1. The activated exonuclease starts the degradation of the strand in a 5'→3' direction. The single-stranded gap is stabilized by RPA. When the mismatch is removed, EXO1 activity is no longer stimulated by MutS, and is actively inhibited by MutL. Pol  $\delta$  loads at the 3' terminus of the original discontinuity, which carries a bound PCNA molecule. This complex fills the gap and DNA ligase I seals the remaining nick to complete the repair process. **b** | Clamps that migrate downstream encounter a PCNA molecule that is bound at the 3' terminus of the strand break. The recruitment and the activation of EXO1 results in the degradation of the region between the original discontinuity and the mismatch through several EXO1-loading events. RFC that is bound at the 5' terminus of the discontinuity prevents degradation in the 5'→3' direction. Once the mismatch is removed and the EXO1 activity is inhibited by bound RPA and MutL, the gap is filled by Pol  $\delta$ . DNA ligase I seals the remaining nick to complete the repair process. Adapted from Jiricny, 2006.

In our study we have focused on *MLH1* gene, which is coding one of the most important components of MMR system. It is often mutated in Lynch syndrome and approximately 15-20 % of sporadic CRCs display malfunction of MMR, including hypermethylation of *MLH1* promoter. *MLH1* gene consists of 19 exons and MLH1 protein has three important domains: ATP-binding, MutS homologues interaction domain and PMS1/PMS2/MLH3 interaction domain (**Figure 7**).



**Figure 7: Diagram of the MLH1 protein in scale.** Numbers inside the blue boxes indicate the exon from which is translated each part of the protein. The three boxes inside represent the ATPase domain, the MutS homologues interaction domain and the *PMS2/MLH3/PMS1* interaction domain; C: Carboxyl-terminal; N: Amino-terminal.

Adapted from <http://atlasgeneticsoncology.org/Genes/MLH1ID149ch3p21.html>.

### 2.2.1. MMR and sporadic colorectal cancer

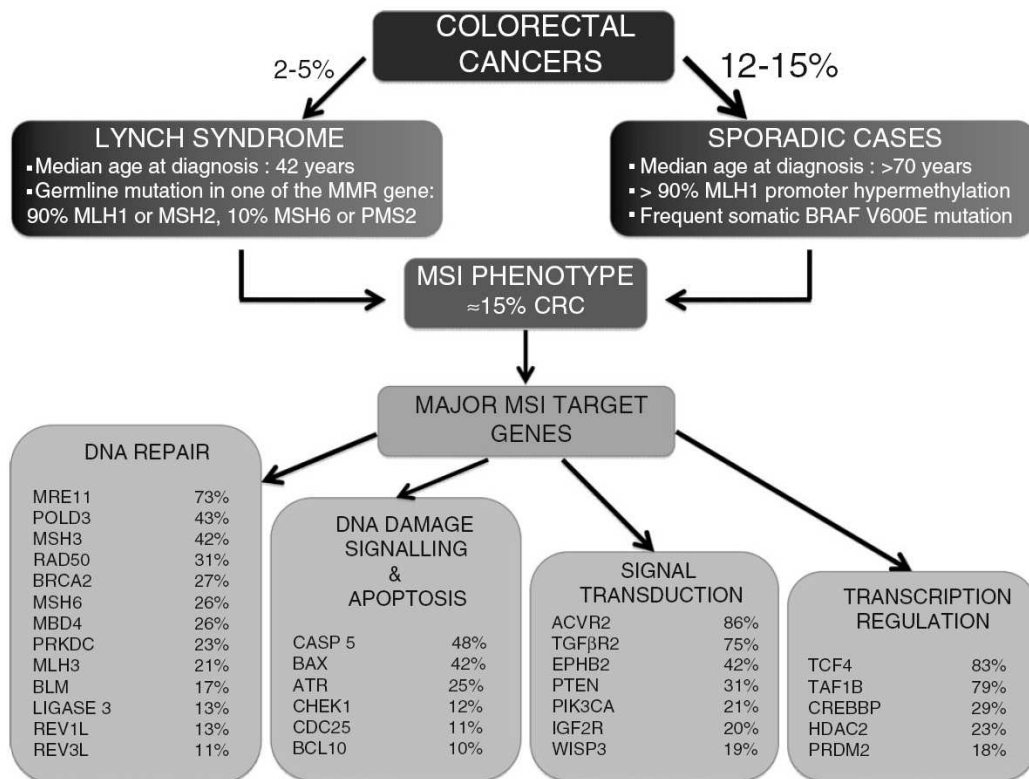
The MSI or pathway is present in approximately 15-20 % of sporadic CRCs and in HNPCC which accounts for approximately 3 % of all CRCs. As it was mentioned MSI tumours are characterized by huge accumulation of mutations in microsatellite sequences. Microsatellites are repeated sequences with motifs such as (A)<sub>n</sub> or (CA)<sub>n</sub> which are present in the genome in large numbers (Ionov et al., 1993). The accumulation of thousands of mutations caused by erroneous activity of DNA polymerase is characteristic for MSI-H (see subchapter 2.2.2.) phenotype and requires the inactivation of MMR genes (Imai and Yamamoto, 2008). Epigenetic changes in *MLH1* and *MSH2* are the most common causes of complete inactivation of MMR, thus defects in other MMR genes play only the partial role in MMR deficiency (Hoeijmakers, 2001).

**Sporadic MSI CRCs:** Typical characteristic of MSI pathway is that sporadic CRCs do not show substantial cytogenetic abnormalities and tumours are not aneuploid (Eshleman et al., 1998). Mutations in the genes *APC*, *KRAS* and *p53* and LOH at chromosomes 5q, 17q and 18q are absent in MSI pathway (Jass et al., 2002).

However, mutations are present in crucial regulatory genes such as *TGFbRII*, *IGFR2R*, *BAX*, *caspase-5*, *b-catenin*, *Tcf-4*, *axin*, *MMP-3*, *E2F-4*, *BCL-10*, *bRAD50* (Morán et al., 2010). Most sporadic CRCs with MSI-H status show the CpG island methylator phenotype (CIMP), caused by DNA hypermethylation (Jass et al., 2002). Sporadic MSI-H tumours arise mainly due to inactivation of *MLH1* gene which results from biallelic promoter hypermethylation rather than somatic mutations or LOH (Kuismanen et al., 2000). According to literature, also microsatellite stable (MSS) forms of CRCs do not exhibit mutations in MMR genes (Zaanan et al., 2011).

MSI-H sporadic tumours display distinct clinical and histopathological features. MSI-H CRCs are more frequent in individuals before the age of 55 or over the age of 70 (Samowitz et al., 2001). They are located more proximally and they are more frequent in women (Gryfe et al., 2000; Raut et al., 2004). Typical features are great depth of invasion, lower stage, local lymphocyte infiltration and lower frequency of metastases (Greenon et al., 2003). Some studies revealed that MSI-H tumours are poorly differentiated and occur in mucinous-cell type tumours (Raut et al., 2004). MSI-H tumours have better prognosis than those with MSI-L or MSS status independently of the molecular processes leading to their development (Samowitz et al., 2001). The reason of better prognosis is that MSI-H cancer display enhanced immunogenic properties mainly by lymphocyte infiltration (Guidoboni et al., 2001).

The main difference in molecular mechanisms (**Figure 8, p. 27**) causing sporadic MSI CRCs and HNPCC is that sporadic MSI-H CRCs are caused mainly by promoter hypermethylation of *MLH1* gene, whereas germ-line mutations in MMR genes followed by „second hit“ are more frequent in HNPCC (Zaanan et al., 2011).



**Figure 8: Main affected genes in MSI pathway and differences between sporadic and hereditary form of CRC with deficient MMR system.** Adapted from Zaanan et al., 2011.

*BRAF* (rapidly accelerated fibrosarcoma, protein type B) mutations are more frequent in sporadic MSI tumours than in HNPCC. HNPCC patients are younger than those with sporadic CRCs (Oliviera et al., 2004).

### 2.2.2. Microsatellite instability (MSI) status and CRC

MSI arises due to inactivation of MMR genes leading to accumulation of mutations in microsatellite sequences. The testing procedure was recommended by National Cancer Institute and it is based on analysis of the tumour and normal tissues using five microsatellite markers. Two mononucleotide markers BAT25 and BAT 26 and three dinucleotide markers D2S123, D5S346 and D17S250 are routinely used (Boland et al., 1998). MSI-low (MSI-L) tumours are defined as having instability in one or two markers. MSI-high (MSI-H) tumours have instability at least in two or more markers. Tumours with no proven instability (20 per cent or less) are termed microsatellite stable (MSS) (Søreide et al., 2006).

The problem in defining MSI status is that not all authors define MSI-L by using the same criteria (regarding both the number and the type of microsatellites analyzed). MSI-H tumours are very successfully analyzed by mononucleotide markers BAT26 and BAT40 (Pawlik et al., 2004). Another problem is that MSI-L tumours have very similar clinicopathological features as MSS tumours thus many authors do not consider them as distinct group of CRCs (Laiho et al., 2002). Typical feature is that LOH at selected chromosomes and *KRAS* mutations are more frequent in MSI-L tumours than in MSI-H (Laiho et al., 2002). Also mutations typical for MSI-H are absent in MSI-L tumours (Imai et al., 2008). Thus it seems that MSI-L CRCs have the same clinicopathological features with MSS tumours but they present different molecular principles. For that reason they are considered as distinct category of CRCs.

### **2.2.3. MSI and CIMP markers of early progression, clinical outcomes and the risk of sporadic CRC**

The promoter region of each gene is the starting position of transcriptional events leading to production of the nascent transcript. For successful transcription, the promoter has to be unmethylated at cytosines and guanines called CpG islands. A common feature of human tumours is that CpG islands are hypermethylated which results in transcription silencing (Issa, 2004). This feature is called CpG island methylator phenotype (CIMP) and cancers with CIMP status represent a clinically and aetiologically distinct group (Toyota et al., 1999). Approximately 80 % of sporadic tumours with MSI-H are associated with the hypermethylation of the promoter region of *MLH1* gene. Samowitz et al. recommended the rules for determining the CIMP status. Tumours with methylation in more than three of genes *cyclin-dependent kinase 2A (CDKN2A)*, *methylated in tumours 1 (MINT1)*, *MINT2*, *MINT31* and *MLH1* are defined as CIMP-high (CIMP-H) and they account for approximately 24-51 % of all CRCs (Samowitz et al., 2005). They are associated with either *BRAF* or *KRAS* mutations, older age, advanced stage and proximal colon location. Recently a new panel of markers has been proposed and it consisted of five loci *calcium channel gene A1G (CACNA1G)*, *insulin growth factor 2 (IGF2)*, *neurogenin 1 (NEUROG1)*, *Runt-related transcription factor X3 (RUNX3)* and

*suppressor of cytokine signaling 1 (SOCS1)*. Thus CIMP-H tumours are defined as those with three to five methylated loci (Weisenberger et al., 2006). Moreover CIMP-low (CIMP-L) tumours were reported with less extensive promoter methylation (Ogino et al., 2006). The molecular mechanisms of CIMP-L tumours are no clear.

MSI and CIMP markers are used as predictive markers in early progression of CRCs. Dinucleotide markers are less sensitive than mononucleotide markers BAT25 and BAT26 for the detection of MSI-H CRCs. Results with dinucleotide markers are prone to an incorrect classification of CRC tumours. Another problem is that ethnic origin of the individuals may produce false-positive results, thus a new panel of quasimorphic MSI markers (BAT25, BAT26, NR-21, NR-22 and NR-24) has been proposed for more sensitive detection of MSI-H tumours (Suraweera et al., 2002). The choice of markers influences the results of testing for MSI or CIMP.

A genetic difference has been noted among colonic versus rectal tumours: *KRAS* mutations are more common in colon than in rectal tumours and the mutational patterns typical for *APC* gene are more common in the rectum (Frattoni et al., 2004). Patients with MSI-H sporadic CRCs have longer and better survival than patients with cancers exhibiting CIN probably due to cytotoxic effects of infiltrated CD8+ lymphocytes (Popat et al., 2005). Many key immunomodulatory genes, such as antigen chaperone molecules, proinflammatory cytokines and cytotoxic mediators are up-regulated in MSI-H cancers (Banerjee et al., 2004). The prognostic advantage of MSI-H tumours has been shown to be most evident in stage II and III (Benatti et al., 2005). However, MSI-positive tumours that express p53 seem to have a more aggressive biology than their p53-negative counterparts (Mori et al., 2004). Questionable is, if 5-fluorouracil (5-FU) chemotherapy is useful for patients with MSI-H. Benefit from this therapy is contradictory because more studies confirmed that stage II or III MSS CRC patients benefit from 5-FU, whereas MSI-H patients do not (Carethers et al., 2004). Thus patients with deficient MMR system cannot display benefits from 5-FU therapy and it should be limited to avoid harmful side effects (Jo and Carethers, 2006). In another study it was revealed that chemotherapy treatment

caused 2-fold increase in mortality among stage III MSI-H patients and 3-fold increase in mortality among stage II MSI-H patients suggesting that chemotherapy has immunosuppressive effects, which ceases enhanced anti-tumour immunity in MSI-H patients (Ribic et al., 2003). Topoisomerase-I inhibitor, irinotecan, is connected with increased survival in patients with MSI CRCs (Fallik et al., 2003). However, generally tumours with MSI are resistant to chemotherapy based on 6-thioguanine, 5-FU, platinum compounds and other agents (Aebi et al., 1997).

### **3. Aims of diploma thesis**

The diploma thesis has the following main aim:

Investigation of somatic events and gene variants in *MLH1* gene in patients with the sporadic form of CRC.

- The secondary aim is to confront the obtained data on *MLH1* gene with those on MSI status (kindly provided by Dr. Schwarzová) and basic clinico-pathological characteristics (kindly provided by Dr. Schneiderová).



## **4. Materials and methods**

### **4.1. Study population**

This study was conducted on 99 DNA samples obtained from 96 patients with sporadic form of CRC. Cases were recruited at the General Teaching Hospital in Prague from 2009 to 2011. Diagnosis was addressed by positive endoscopic results for malignancy, and subsequently histologically confirmed as carcinomas of colon or rectum. From each patient we obtained samples of healthy colon mucosa, peripheral blood and tumour tissue. DNA samples bearing known germ-line variants in exons 3, 5, 8, 12 and 14 in the *MLH1* gene were used as positive controls. These samples were obtained from patients with the diagnosis of Lynch syndrome, recruited from the General Teaching Hospital in Prague and from the Thomayer Hospital in Prague. All cases included in the study have been matched for age and sex. Within a study a structured questionnaire was used to collect information from study subjects about lifestyle habits (smoking, drinking, diet, etc.), and family/personal history of cancer.

### **4.2. Ethical approval**

An informed written consent was signed by all participants and the study design (GACR P304/12/1585 Molecular characteristics of DNA repair in tumour tissues of CRC patients) was approved by the Ethical Committee of the First Faculty of Medicine, Charles University and General Teaching Hospital in Prague, Czech Republic.

### **4.3. DNA processing**

DNA from tumour tissues and healthy mucosa tissues was isolated using QIAamp DNA Mini Kit (Qiagen, Germany) according to manufacturer instructions. DNA from peripheral blood of the same patients was isolated using QIAcube (Qiagen, Germany) machine. Above procedures were done by technicians (Martina Řezníčková, Michaela Krylová and Monika Škodová) at the Institute of Biology and Medical Genetics. DNA with germ-line mutation in exon 3, used as positive control, has been obtained by Dr. Martina Langová, Thomayer Hospital in Prague. All DNA samples were diluted to final concentration of 10 ng/μl in a total volume of 100 μl.

The concentration of DNA was measured spectrophotometrically on the NanoPhotometer® (Implen, Germany). The HRM was performed in a reaction mix containing 20 ng of genomic DNA (i.e. 2 µl of DNA solution); the total volume of HRM reaction was 10 µl. For PCR reactions of total volume of 25 µl, 1 µl of DNA template were used. For sequencing PCR of total volume of 5 µl; 0.3 µl of DNA template were used.

#### 4.4. Determination of MSI status

All DNA samples were analyzed for MSI using pentaplex PCR according to Buhard et al., 2006; the data on MSI have kindly been provided for this study by Dr. Schwarzová. Five markers were coamplified in multiplex PCR (programme: denaturation at 95°C for 5 minutes; 35 cycles of amplification, denaturation at 95°C for 30 seconds, annealing at 55°C for 30 seconds and extension at 72°C for 30 seconds, followed by extension at 72°C for 5 minutes; **Table 2**).

Marker	Gene	GenBank No.	Localization	Primer sequences	Size A (bp)	Size B (bp)
NR-27	Inhibitor of apoptosis Protein-1	AF070674	5' UTR	AACCATGCTTGCAAACCA CT CGATAATACTAGCAATGA CC	90	87
NR-21	<i>SLC7A8</i>	XM_033393	5' UTR	GAGTCGCTGGCACAGTTC TA CTGGTCACTCGCGTTTAC AA	110	107
NR-24	Zinc finger2	X60152	3' UTR	GCTGAATTTTACCTCCTG AC ATTGTGCCATTGCATTCC AA	129	126
BAT-25	<i>c-kit</i>	X69313	Intron 16	TACCAGGTGGCAAAGGG CA TCTGCATTTTAACTATGGC TC	152	148
BAT-26	<i>MSH2</i>	AY601851	Intron 5	CTGCGGTAATCAAGTTTT TAG AACCATTCAACATTTTAA CCC	182	179

**Table 2: MSI markers for pentaplex PCR.** MSI status was determined in all DNA samples using five different markers. In each case, the antisense primer was labeled with a fluorescent dye: FAM for BAT-26 and NR-21, HEX for BAT-25 and NR-27, and NED for NR-24. The indicated sizes are (a) calculated from GenBank (National Institutes of Health, Bethesda, MD), or (b) deduced from the more frequent allele obtained experimentally in the worldwide germ-line DNA analysis. Abbreviation: bp, base pairs. Modified from Buhard et al., 2006.

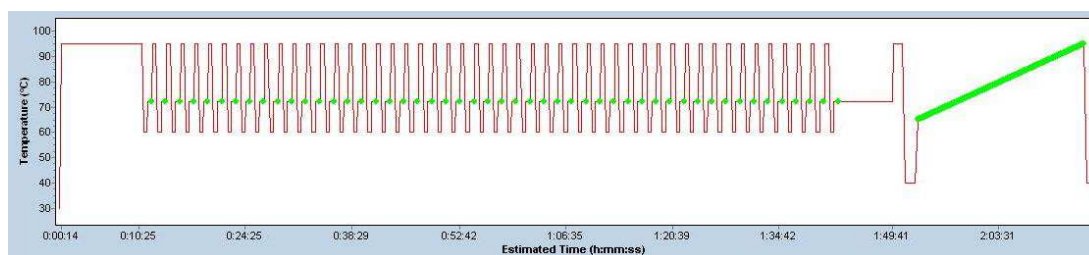
All amplified samples were analyzed on ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, USA). Following the PCR amplification of tumour DNA, MSI was defined by the presence of novel peaks that were not present in the referent DNA.

#### **4.5. Mutational analysis of *MLH1* gene – HRM, PCR**

High resolution melting (HRM) is a cost-effective and rapid method for detecting variations of bases in DNA samples. It is based on precise monitoring of the change in fluorescence caused by the release of an intercalating DNA dye from a DNA duplex as it denatures at high temperatures (Rouleau et al., 2009). We used LightCycler® 480 and LightCycler® 480 Master kit, both developed by Roche Diagnostics, Germany. The HRM analysis is based on the following processes:

1. First, the region of interest within the DNA sequence is amplified using the PCR performed in the presence of the fluorophor which is intercalated into double strand DNA (dsDNA). Concentration of DNA is measured after each PCR cycle. As the amplicon concentration in the reaction tube increases, the fluorescence exhibited by the double stranded amplified product also increases.
2. After all PCR cycles, dsDNA is heated to its melting temperature producing single strand DNA (ssDNA) and then rapidly cooled to 40°C. Double-strand DNA heteroduplexes are produced.
3. HRM analysis: the amplicons of DNA are heated gradually from around 50°C up to around 95°C. At a melting point DNA denatures and the fluorescence fades away.
4. The HRM machine uses a camera that measures the amount of fluorescence emitted as DNA denatures. The data from this analysis are plotted in a graph known as a Melting curve, showing the level of fluorescence vs. the temperature. If one of the patients has a mutation in the DNA region we have amplified, then this will alter the temperature at which the DNA strands melt apart. The melting curve representing this patient will have different shape.

HRM analysis was carried out using the following programme (**Figure 9, Table 3**):



**Figure 9: The overview of HRM analysis programme;** estimated time on the x axis and temperature in °C on the y axis. The proper HRM analysis is represented by green line on the right.

Phase of HRM	Temperature (°C)	Time (min)
Activation of DNA polymerase	95	10:00
Amplification	95	00:15
	60	00:15
	72	00:25
Final elongation	72	07:00
HRM analysis	95	10:00
	40	01:00
	65	00:01
	65 → 95 (0.02°C/sec)	continuous
Cooling	40	00:10

**Table 3: The programme of HRM analysis used for *MLH1* gene.** All exons of *MLH1* gene except of exons 12 and 18, were analyzed using this programme.

In this study we have analyzed all 19 exons within *MLH1* gene. Exon 12 has been divided into two overlapping regions (12-A; 12-B) because the optimal length of PCR product for HRM is up to 400 bps. The length of the exon 12 is 371 bps, but PCR product length is 550 bps. In general, all primers used in this study have been adapted according to Rouleau et al., 2009 (**Table 4, p. 36**). Primers were purchased

from Sigma Aldrich, USA. For exons 12 and 18 new primers have been designed using Primer3 software due to non-specific amplification of exon 18, and unsuccessful PCR amplification of exon 12-A, respectively. These primers were purchased from Generi Biotech, Czech Republic (**Table 5**).

Amplicons <i>MLH1</i>	Forward primer	Reverse primer	Amplicon size
1	AGGCACTGAGGTGATTGGC	CCGTAAAGTCGTAGCCCTTA	237
2	AGTTTGTATCATTGCTTGGTCAT	TCCAGAACAGAGAAAGTCCCTGACT	250
3	TCAAGAAAATGGGAATTCAAAGAGAT	TTCATTAAGTTTGCTCAGATTGTCAT	323
4	CTTTGGTGAGGTGACAGTGG	GATTACTCTGAGACCTAGGC	217
5	CTCTTTTCCCCTTGGGATTAGTATC	CAATATTTATACAAACAAAGCTTCAACA	195
6	TAAGACCTCGCTTTTGCC	GCTCAGCAACTGTTCAATGTATGAGC	252
7	TAAAAGGGGGCTCTGACATCTAGTG	CATAACCTTATCTCCACCAGCAAACCT	203
8	GCTGGTGGAGATAAGGTTATGATGT	CTTGAAAGGTTCCAAAATAATGTGAT	255
9	AGTTAGTTTATGGAAGGAACCTTGT	CCCTGTGGGTGTTTCTGTGA	268
10	CATGACTTTGTGTGAATGTACACC	GAGGAGAGCCTGATAGAACATCTG	239
11	TCTCCCCCTCCCACTATCTAAGGTA	GCCCCAGAGAAGTAGCTGGATGA	241
12-1	TACTTCTTATTCTGAGTCTCTCCACTA	ATATCTGTCTTATCCTCTGTGACAAT	295
12-2	TGATAAGGTCTATGCCACCAGAT	AGGCAGAGAGAAGATGCAAGTGAT	387
13	CATTACAGCTCTGTAGAACCAGCA	ATTTCCAAACCTTGGCAGTTGA	335
14	TGGTGTCTCTAGTCTCTGG	CCATTGTTGTAGTAGCTCTGC	255
15	TTTGTCCCACTGGTTGTATCT	TGGAGAGCTACTATTTTCAGAAACGA	200
16	CATTGGATGCTCCGTTAAAGC	CACCCGGCTGGAAATTTTATTTG	276
17	CCAGAGTGGCAGATAGGAGCACAC	CCCTTTCCCTCCAGCACACAT	263
18	TGTGATCTCCGTTTAGAATGAGAATGT	AAAGATTGTATGAGGTCCTGTCTAGT	243
19	GCAACAGGGAGGCTTATGA	CAGAGAAAGAAGAACACATCC	250

**Table 4: List of primers used for HRM analysis.** All primers used in HRM analysis are summarized in the following table. Adapted from Rouleau et al., 2009.

Exon	Forward primer	Reverse primer	Length (bps)
12-A	TACTTCTTATTCTGAGTCTCTCCACTA	ATATCTGTCTTATCCTCTGTGACAAT	295
18	TAAGTAGTCTGTGATCTCCG	ATGTATGAGGTCCTGTCC	247
18	GTCTGTGATCTCCGTTTAGAATGA	TCCTAAAGATTGTATGAGGTCCTGT	250
18	CTGAAGGCCTGCTCTCATCCCC	GTACCTGTGGCCTGAGAGGGT	301

**Table 5: Primers designed by Primer3 software for exons 12-A and 18**

Each HRM reaction was performed in a single run in a reaction mix containing 20 ng of the genomic DNA, 5 µl of LightCycler® 480 High Resolution Melting Master master mix containing ResoLight® dye, 3 mM MgCl<sub>2</sub>, and 0.4 µM of each primer with PCR-grade water, adjusted to a total volume of 10 µl per well. For

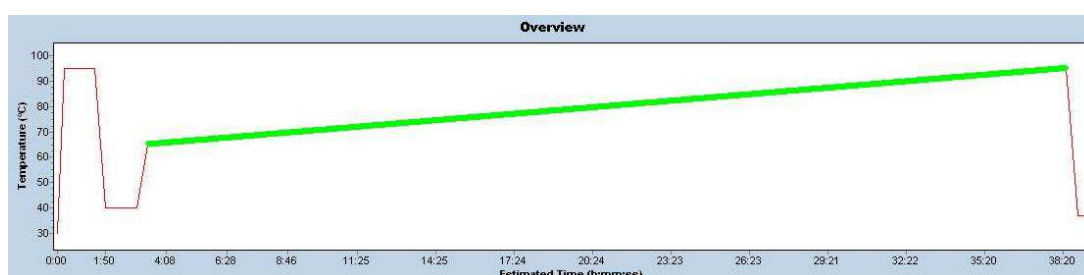
each HRM reaction 2 µl of DNA sample were used. For this purpose we used Eppendorf epMotion® 5075 Automated Pipetting System (Eppendorf, Germany).

Exons 12-A and 18 were analyzed by automated sequencing because of unsuccessful optimization of HRM reaction.

Positive control is used for comparison of melting curves. If we do not use the homozygous carrier of gene variant as positive control, it is not possible to detect variant homozygotes in HRM. However, heterozygous carriers of gene variant are easily detected by HRM. For this reason, PCR product of normal DNA (not bearing SNPs) sample was prepared. Five µl of HRM product and 5 µl of normal DNA product, amplified by PCR, were mixed together. It allowed us mimic the heterozygous carriers of gene variant, which were analyzed by LightCycler® 480. This analysis was based on the following programme (**Figure 10, Table 6**):

Phase of HRM	Temperature (°C)	Time (min)
HRM analysis	95	01:00
	40	01:00
	65	00:01
	65 → 95 (0.02°C/sec)	continuous
Cooling	37	10:00

**Table 6: Programme used in HRM analysis of homozygous carriers of gene variant.**



**Figure 10: The overview of HRM.** Estimated time is shown on the x axis and temperature on the y axis.

Normal DNA samples used for this second step were prepared by PCR performed in the total volume of 25 µl. The reaction master mix contained following chemicals: 2.5 µl of 10× PCR buffer II (Applied Biosystems, USA); 1.7 mM MgCl<sub>2</sub>

solution (Applied Biosystems, USA); 20  $\mu$ M dNTP Mix (Fermentas, USA); 0.75 U of AmpliTaq Gold<sup>®</sup> DNA polymerase (Applied Biosystems, USA); 0.2 mM of each primer (Sigma Aldrich, USA; Generi Biotech, Czech republic); 10 ng of DNA template and 18.4  $\mu$ l of H<sub>2</sub>O (Braun Medical, USA). The PCR conditions are summarized in **Table 7**. Some DNA samples had to be sequenced by automated sequencing. For that purpose we used DNA templates prepared on HRM or by standard PCR. These PCR reactions were run in the total volume of 10  $\mu$ l. Concentrations of all chemicals were adjusted according to reaction volume of 10  $\mu$ l.

Phase of PCR	Temperature (°C)	Time (min)
Activation of DNA polymerase	95	05:00
Amplification	95	00:30
	60	00:30
	72	00:45
Cooling	4	$\infty$

**Table 7: The programme of PCR used for amplification of normal DNA samples** (used in HRM analysis of homozygous carriers of gene variant).

All chemicals were vortexed before use by GV Lab Vortex (Gilson, France). All manipulation with DNA was done on the cooling block. PCR was run on the MyCycler thermo cycler (Bio-Rad, USA) and on the Vapo.protect cycler (Eppendorf, Germany). For control of PCR amplification, agarose electrophoresis was used. For each PCR product 5  $\mu$ l of product was mixed together with 3  $\mu$ l of concentrated sample buffer. This mix was loaded onto 2% agarose gel (Serva, Germany). Power Pac Basic or Power Pac 300 was used as power sources (Bio-Rad, USA). Electrophoresis was run on a constant current of 160 V. We have used Gene Ruler<sup>™</sup> 100 bp Plus DNA ladder (Fermentas, USA). Each electrophoresis was run approximately 45 minutes. DNA was visualized on LKB 2011 MacroVue UV transilluminator (Bromma, Sweden). Detection is based on the fluorescent dye GelRed (Biotium Inc, USA), which is added into agarose gel. DNA in some PCRs was insufficiently amplified or more unspecific products were amplified. For those reasons we used 0.28 mM DMSO (MP Biomedicals, USA) and Surveyor Nuclease Enhancer S (Transgenomic, USA).

#### 4.6. Mutational analysis of *MLH1* gene – automated sequencing

Samples analyzed by automated sequencing were first amplified by sequencing PCR. Samples were prepared under the following conditions: 2 µl of Big Dye<sup>®</sup> Terminator v1.1 Cycle Sequencing kit (Applied Biosystems, USA); 2.4 µl of H<sub>2</sub>O; 0.3 µl of DNA template and 0.3 µl of forward or reverse primer depending on the analyzed DNA strand. All samples were prepared in the strip on the cooling block. Strips were centrifuged before PCR, using GMC Lab minicentrifuge (Gilson, France). Programme for sequencing PCR is shown in **Table 8**:

Phase of PCR	Temperature (°C)	Time (min)
Activation of DNA polymerase	95	03:00
Amplification	96	00:20
	50	00:10
	60	04:00
Final elongation (1 cycle)	60	01:00
Cooling	4	∞

**Table 8: The programme of sequencing PCR.**

After PCR, all samples were purified according to following protocol:

1. A half µl of sodium acetate (Penta chemicals, Czech Republic) and 12.5 µl of 96% ethanol (Penta chemicals, Czech Republic) were added to each sample. Samples were left at room temperature for 10 to 15 min.
2. Samples were centrifuged for 30 minutes (11 800 rpm) using Centrifuge 5430 (Eppendorf, Germany) or 23 minutes (14 500 rpm) using miniSpin Plus (Eppendorf, Germany).
3. Supernatant was removed by rapid hand movement. Then 60 µl of 70% ethanol (Penta chemicals, Czech Republic) were added to each sample.
4. Samples were centrifuged for 10 minutes using Centrifuge 5430 (11 800 rpm, Eppendorf, Germany) or miniSpin Plus (14 500 rpm, Eppendorf, Germany). Supernatant was removed after centrifugation and 60 µl of 70% ethanol were



added to the pellet. Samples were centrifuged for 10 minutes again.

5. Supernatant was removed and pellet was dried at the heating block Infinigen Pro III for 10 minutes at 40°C (Infinigen Biotech, USA).
6. Then 30 µl of formamide was added to each dry pellet.
7. Samples were denatured at the heating block Infinigen Pro III for 5 minutes at 95°C.
8. Purified samples were put to cooling block prepared for automated sequencing.

All DNA samples were sequenced by automated sequencing (Applied Biosystems® 3130 Genetic Analyzer; Applied Biosystems, USA) at the Institute of Biology and Medical Genetics.

#### **4.7. Data processing**

Primers for exons 12-A and 18 were designed using Primer3 programme. Reference sequence of MLH1 gene NG\_007109.1 from NCBI database was used. Primer3: <http://frodo.wi.mit.edu/primer3/>

Steve Rozen and Helen J. Skaletsky (2000) Primer3 on the WWW

NCBI reference sequence of MLH1: NG\_007109.1

[http://www.ncbi.nlm.nih.gov/nuccore/NG\\_007109.1?from=4863&to=62359&report=genbank](http://www.ncbi.nlm.nih.gov/nuccore/NG_007109.1?from=4863&to=62359&report=genbank)

The control of electrophoretograms from sequencing was done manually and by BioEdit software. All variants in *MLH1* gene were presented according to nomenclature published by Human Genome Variation Society.

Nomenclature: <http://www.hgvs.org/mutnomen/>

Programme BioEdit: <http://www.mbio.ncsu.edu/bioedit/bioedit.html>

All variants found in this study were compared with Human Genome Mutation Database (HGMD), Mooney Lab – MutDB, InSiGHT database and in the Mismatch Repair Genes Variant Database. SNPs were compared with GeneCards and NCBI databases.

HGMD: <http://www.hgmd.org/>

Mooney Lab – MutDB: <http://mutdb.org/>

InSiGHT: <http://www.insight-group.org/>

MMR Genes Variant Database: <http://www.med.mun.ca/mmrvariants/>

GeneCards: <http://www.genecards.org/>

NCBI SNP: <http://www.ncbi.nlm.nih.gov/snp/>

## 5. Results

### 5.1. Clinical characteristics of CRC patients

In this study **Table 9 (p. 43)** summarizes information on age, sex and year of birth for each patient, whose DNA was analyzed. Clinical and pathological characteristics for each CRC patient, involving stage of the malignancy, family history of cancer, body mass index, smoking status, stress factors, vegetarian diet, education and living habitat, were kindly provided by Dr. Schneiderová. Above characteristics along with MSI status (kindly provided by Dr. Schwarzová) are presented in **Table 10, p. 44**.

No. of sample	Year of birth	Age	Sex
17	1929	83	Z
19	1927	85	M
21	1924	88	M
23	1957	55	Z
25	1943	69	M
27	1941	71	Z
29	1948	64	Z
31	1937	75	M
33	1950	62	M
35	1935	77	Z
37	1926	86	Z
39	1943	69	Z
41	1951	61	M
43	1936	76	M
45	1936	76	Z
47	1947	65	M
49	1950	62	M
51	1929	83	Z
53	1935	77	M
55	1927	85	M
57	1942	70	M
59	1956	56	M
61	1944	68	Z
63	1941	71	M
65	1925	87	Z
67	1949	63	M
69	1934	78	Z
71	1941	71	M
73	1925	87	M
75	1940	72	Z
77	1935	77	M
79	1932	80	M
81	1939	73	Z
83	1945	67	M
85	1949	63	M
87	1927	85	Z
89	1972	40	Z
91	1942	70	M
93	1934	78	Z
95	1958	54	M
97	1959	53	M
99	1951	61	Z
101	1945	67	Z
105	1953	59	Z
107	1927	85	M
109	1935	77	Z
111	1933	79	M
113	1942	70	M
115	1946	66	Z
117	1947	65	Z
119	1929	83	Z
121	1943	69	Z

123	1945	67	M	171	1943	69	Z
125	1931	81	M	173	1935	77	Z
127	1925	87	M	175	1943	69	M
129	1930	82	M	177	1920	92	M
131	1938	74	Z	179	1957	55	M
133	1943	69	M	181	1942	70	M
135	1954	58	M	183	1966	46	M
137	1950	62	M	185	1939	73	M
139	1943	69	M	187	1936	76	Z
141	1931	81	M	189	1941	71	Z
143	1928	84	M	191	1930	82	Z
145	1925	87	M	193	1939	73	Z
147	1949	63	M	195	1939	73	Z
149	1941	71	M	197	1939	73	Z
151	1944	68	Z	199	1939	73	Z
153	1941	71	M	201	1936	76	Z
155	1948	64	Z	203	1953	59	Z
157	1953	59	Z	205	1958	54	Z
159	1953	59	Z	207	1931	81	M
161	1954	58	M	209	1930	82	M
163	1948	64	M	211	1929	83	Z
165	1938	74	Z	213	1941	71	M
167	1939	73	M	215	1946	66	Z
169	1944	68	M	217	1942	70	M

**Table 9: List of DNA samples used in the analysis of somatic events in *MLH1* gene providing information on year of birth, age and sex.**

Characteristic	Category	CRC patients
Gender		<b>n=96</b>
	Female	41
	Male	55
Age		<b>n=96</b>
	Mean±SD	69.53 ± 7.87
	Median (quartiles)	72 (38-91)
	Range	38-91
Diagnosis		<b>n=99</b>
	Colon	80 (80.8%)
	Rectum	19 (19.2%)
NM		<b>n=99</b>
	I	24 (24.2%)
	II	28 (28.3%)
	III	28 (28.3%)
	IV	19 (19.2%)

Smoking status		<b>n=96</b>
	Non-smokers	58 (60.4%)
	Smokers	12 (12.5%)
	Ex-smokers	26 (27.1%)
Body Mass Index		<b>n=96</b>
	Mean±SD	24.7 ± 3.6
	< 18.5	2 (2.1%)
	18.5 - 24.9	43 (44.8%)
	25.0 - 29.9	42 (43.8%)
	30.0 - 40.0	9 (9.4%)
Family history of cancer		<b>n=96</b>
	Negative	40 (42%)
	Positive	56 (58%)
CRC in family		<b>n=96</b>
	No	76 (79.2%)
	Yes	20 (20.8%)
Stress factors		<b>n=96</b>
	No	42 (43.8%)
	Yes	54 (56.2%)
Vegetarian diet		<b>n=96</b>
	No	96 (100%)
	Yes	0 (0%)
Living habitat		<b>n=96</b>
	City	53 (55.2%)
	City and country	19 (19.8%)
	Country	24 (25.0%)
Education		<b>n=96</b>
	Basic	54 (56.3%)
	High school	26 (27.0%)
	University	16 (16.7%)
MSI status		<b>n=96</b>
	MSS	87 (90.6%)
	MSI-Low	0 (0%)
	MSI-High	9 (9.4%)

**Table 10: Characteristics of the study population**

## **5.2. Mutations in *MLH1* gene**

In summary we have analyzed 99 DNA samples from 96 patients with sporadic CRC. We have found two new variants in *MLH1* gene, one of them represent somatic mutation and the second one may represent polymorphism or germ-line mutation. We have also found one variant previously described in HNPCC

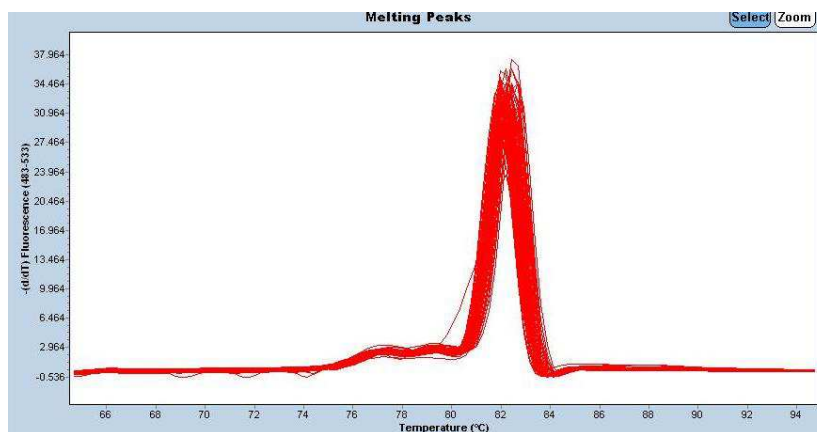
and 5 known gene variants; two of them were located in intronic region of *MLH1* gene. MSI status has been analyzed in all samples; 9 of them displayed MSI-H status, the rest of them were considered as MSS. All variants found in *MLH1* gene are summarized in the **Table 11, p. 57**. Gene variants found in samples displaying MSI-H are summarized in **Table 12, p. 58**.

#### A. New variants in *MLH1* gene

We found a new variant c. 204 C>G, p. Ile68Met in exon 2 in one DNA sample. **Figure 11 A** describes the spectra from HRM so called „difference plot“, in which we can see curves with different shape. These curves represent patients with different variants of any amplified exon of *MLH1* gene. Heterozygous carrier of the gene variant c. 204 C>G is represented by red curve. On the **Figure 11 B** (p. 46) we can see „T<sub>m</sub> calling“ analysis, which describes the specificity of PCR preceding the HRM analysis. If there are unspecific PCR products, melting peaks have different shape according to changes in melting temperature. On the last figure there is an electrophoretogram obtained from automated sequencing containing variant c. 204 C>G compared with wild-type allele of exon 2.

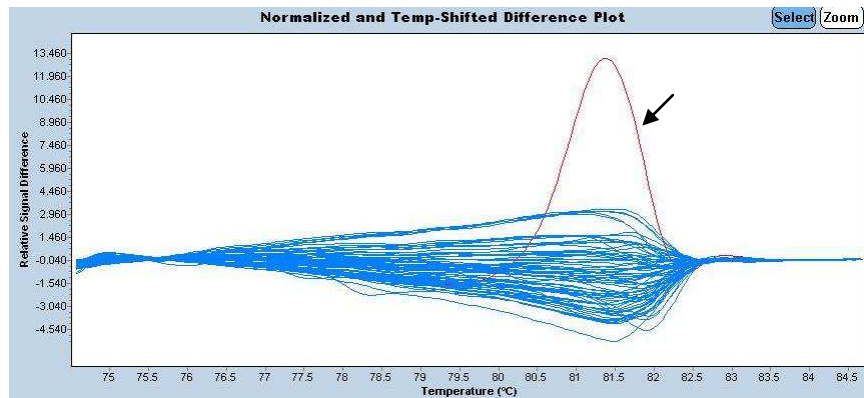
**Figure 11:**

A.

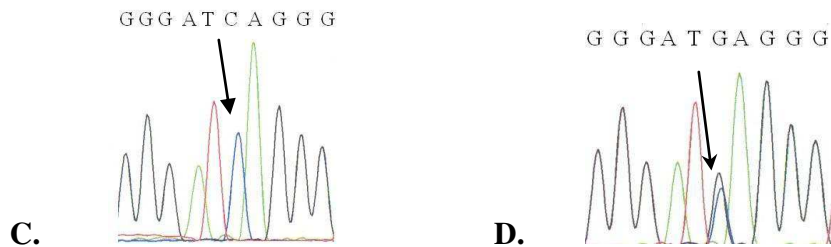


- A. “T<sub>m</sub> calling” analysis of the exon 2 of *MLH1* gene. All peaks have approximately the same shape, thus there is only one PCR product in this analysis. Temperature is displayed on the x axis and the rate of fluorescence on the y axis.

**B.**



**B. Genetic analysis of the exon 2 of *MLH1* gene.** All curves have the same shape except of one representing sample bearing variant c. 204 C>G. The relative signal of fluorescence represented by y axis is changing due to rising temperature and presence of mutations in *MLH1* gene.



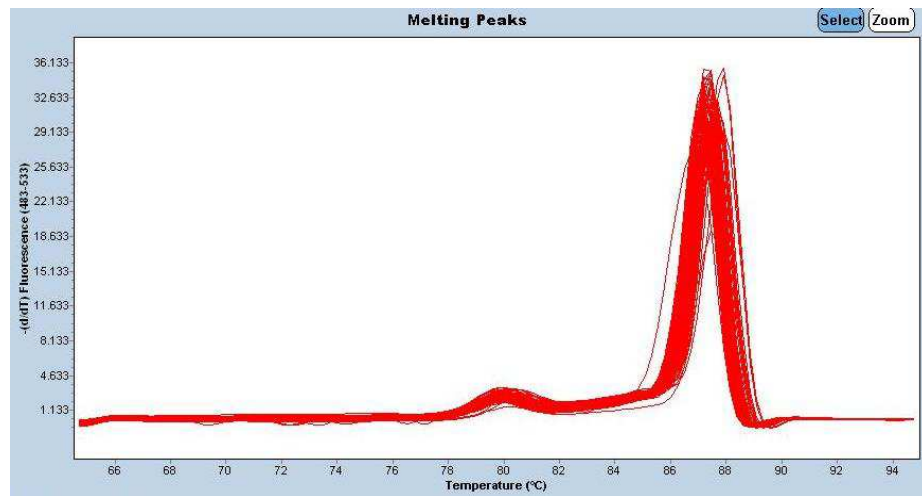
**C.** The result of sequencing of exon 2 of *MLH1* gene without variant c. 204 C>G (wild-type).

**D.** The result of sequencing of exon 2 of *MLH1* gene representing heterozygous carrier of gene variant c. 204 C>G – two peaks at position c. 204.

The second new variant, c. 973 C>T, p. Arg325Trp (**Figure 12, p. 47**), was found in exon 11 in one DNA sample and was assigned as somatic mutation of one allele. This result was confirmed by analysis of DNA sample from peripheral white blood cells, where no such change was observed.

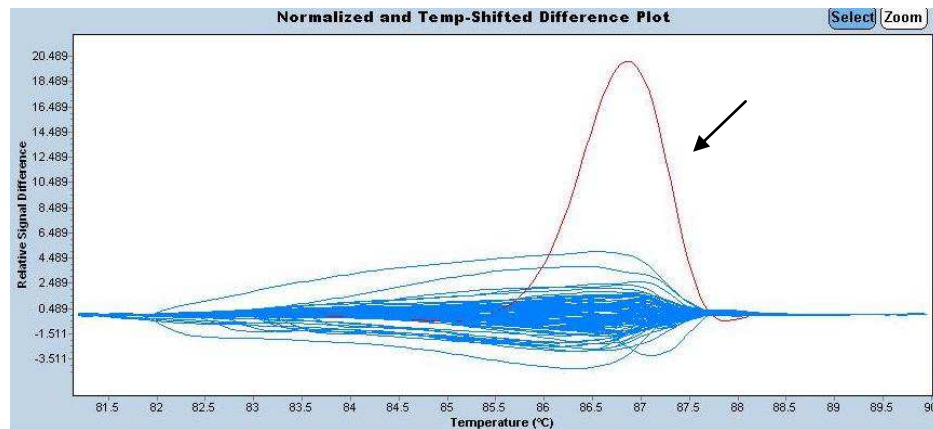
**Figure 12:**

**A.**



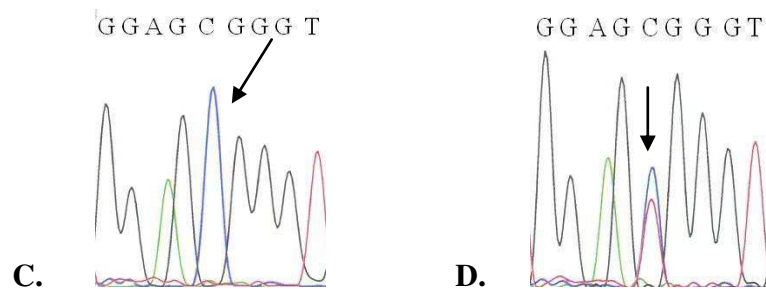
**A. “T<sub>m</sub> calling” analysis of the exon 11 of *MLH1* gene.** All peaks have approximately the same shape, thus there is only one PCR product in this analysis. Temperature is displayed on the x axis and the rate of fluorescence on the y axis.

**B.**



**B. Genetic analysis of the exon 11 of *MLH1* gene.** All curves have the same shape except of one curve representing sample bearing somatic variant c. 973 C>T. The relative signal of fluorescence represented by y axis is changing due to rising temperature and presence of mutations in *MLH1* gene.





**C.** The result of sequencing of exon 11 of *MLH1* gene without variant c. 973 C>T (wild-type).

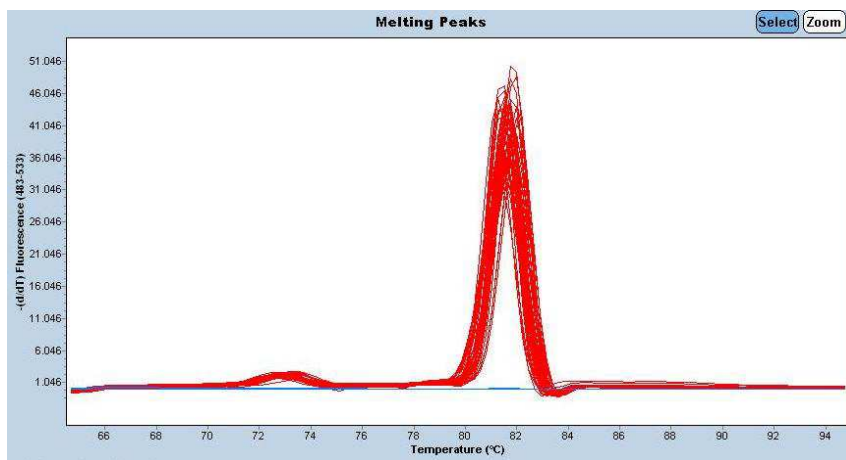
**D.** The result of sequencing of exon 11 of *MLH1* gene representing heterozygous carrier of gene variant c. 973 C>T – two peaks at position c. 973.

### B. Gene variants found in *MLH1* gene with amino acid change

We found polymorphism c. 655 A>G, p. Ile219Val in exon 8 in 50 samples, 16 of them had homozygous combination of alleles (**Figure 13**). The remaining 34 samples had heterozygous combination of alleles.

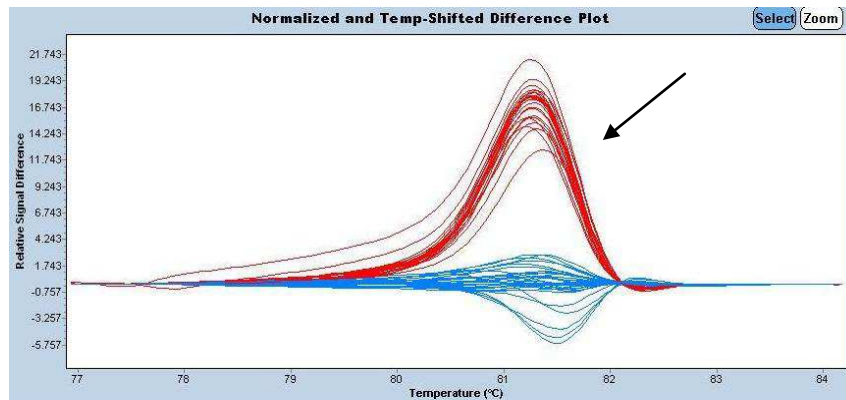
**Figure 13:**

**A.**

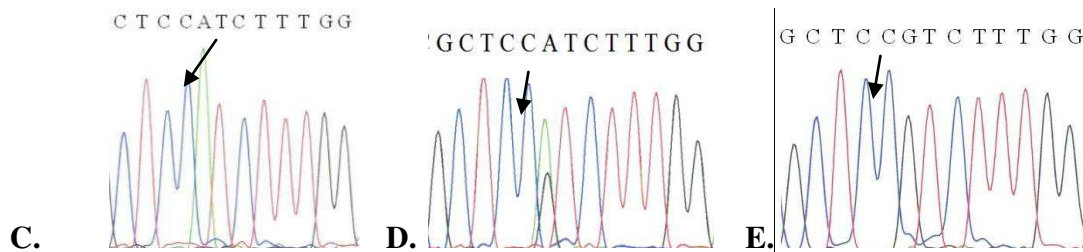


**A.** “T<sub>m</sub> calling” analysis of the exon 8 of *MLH1* gene. All peaks have approximately the same shape, thus there is only one PCR product in this analysis. Temperature is displayed on the x axis and the rate of fluorescence on the y axis.

**B.**



**B. Genetic analysis of the exon 8 of *MLH1* gene.** All curves have the same shape except those representing samples bearing variants c. 655 A>G. The relative signal of fluorescence represented by y axis is changing due to rising temperature and presence of mutations in *MLH1* gene.



**C.** The result of sequencing of exon 8 of *MLH1* gene without variants c. 655 A>G (wild-type).

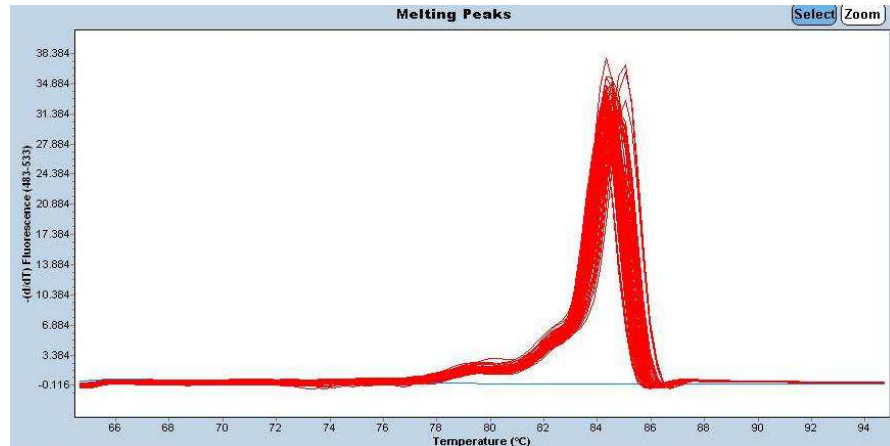
**D.** The result of sequencing of exon 8 of *MLH1* gene representing heterozygous carrier of gene variant c. 655 A>G – two peaks at position c. 655.

**E.** The result of sequencing of exon 8 of *MLH1* gene representing homozygous carrier of gene variant c. 655 A>G.

Germ-line variant c. 1733 A>G, p. Glu578Gly was found in exon 16 in one sample as heterozygous combination of alleles (**Figure 14**).

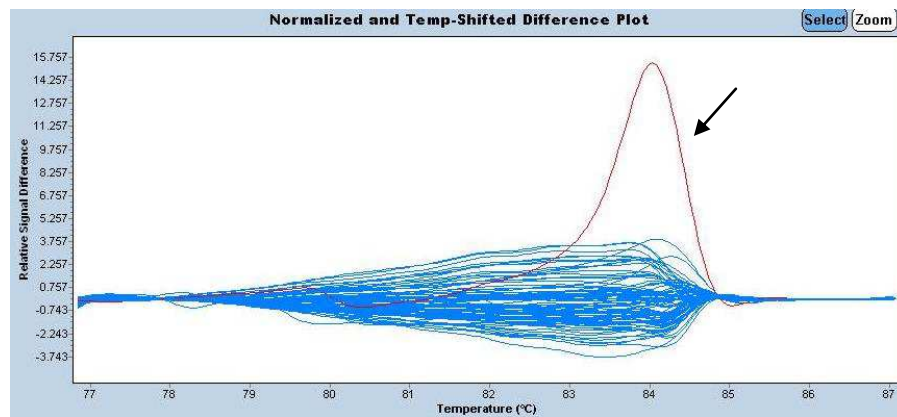
**Figure 14:**

**A.**

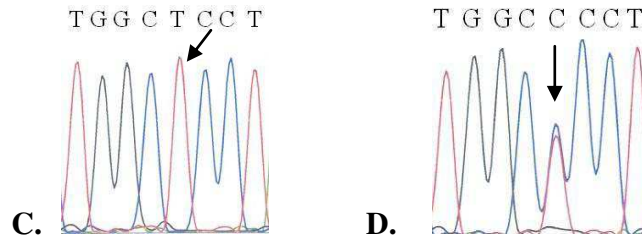


**A. “T<sub>m</sub> calling” analysis of the exon 16 of *MLH1* gene.** All peaks have approximately the same shape, thus there is only one PCR product in this analysis. Temperature is displayed on the x axis and the rate of fluorescence on the y axis.

**B.**



**B. Genetic analysis of the exon 16 of *MLH1* gene.** All curves have the same shape except of one curve representing sample bearing variant c. 1733 A>G. The relative signal of fluorescence represented by y axis is changing due to rising temperature and presence of mutations in *MLH1* gene.



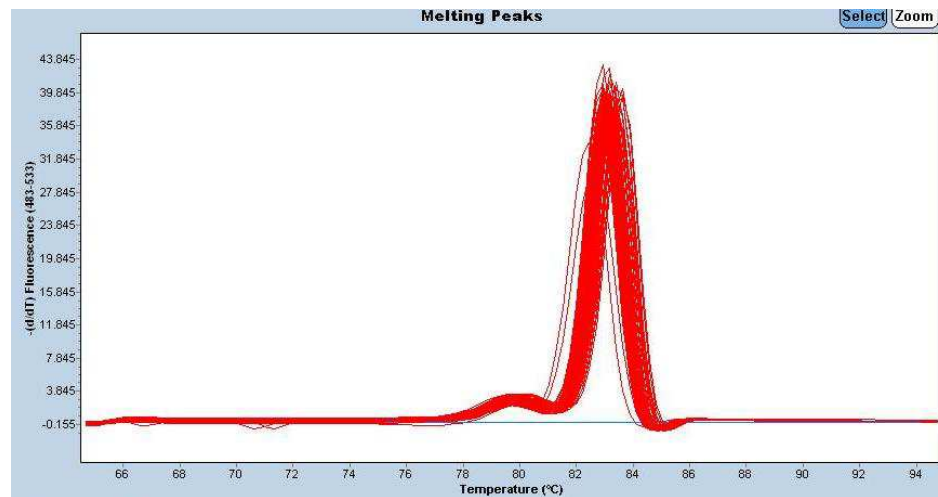
**C.** The result of sequencing of exon 16 of *MLH1* gene without variant c. 1733 A>G (wild-type, REVERSE strand).

**D.** The result of sequencing of exon 16 of *MLH1* gene representing heterozygous carrier of gene variant c. 1733 A>G – two peaks at position c. 1733 (REVERSE strand).

We found polymorphism c. 2146 G>A, p. Val716Met in exon 19 in 2 samples as heterozygous combination of alleles (**Figure 15**).

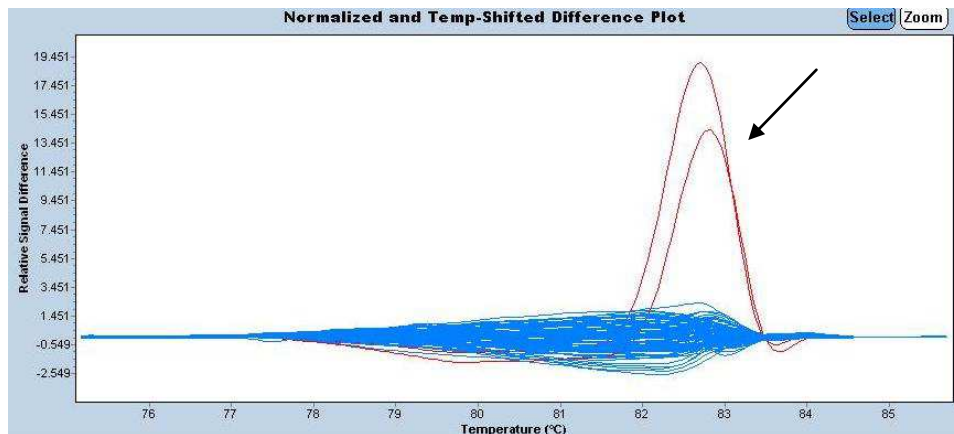
**Figure 15:**

**A.**

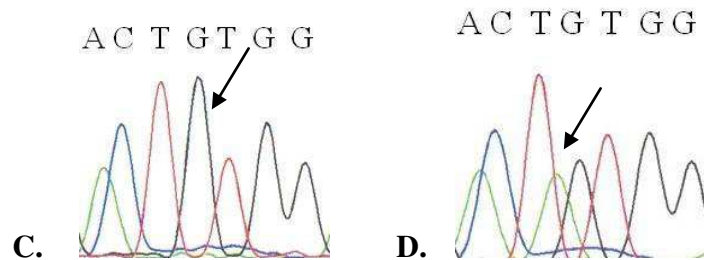


**A.** “T<sub>m</sub> calling” analysis of the exon 19 of *MLH1* gene. All peaks have approximately the same shape, thus there is only one PCR product in this analysis. Temperature is displayed on the x axis and the rate of fluorescence on the y axis.

**B.**



**B. Genetic analysis of the exon 19 of *MLH1* gene.** All curves have the same shape except of two curves representing samples bearing variants c. 2146 G>A. The relative signal of fluorescence represented by y axis is changing due to rising temperature and presence of mutations in *MLH1* gene.



**C.** The result of sequencing of exon 19 of *MLH1* gene without variants c. 2146 G>A (wild-type).

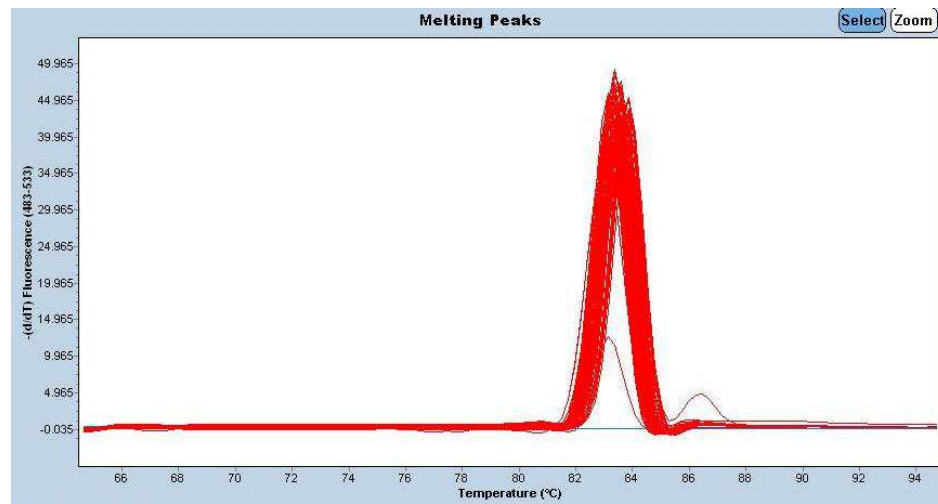
**D.** The result of sequencing of exon 1 of *MLH1* gene representing heterozygous carrier of gene variant c. 2146 G>A – two peaks at position c. 2146.

### **C. Gene variants in *MLH1* gene with no change of amino acid**

Gene variant c. 1959 G>T, p. Leu653Leu was found in exon 17 in 5 DNA samples as the heterozygous combination of alleles (**Figure 16, p. 53**).

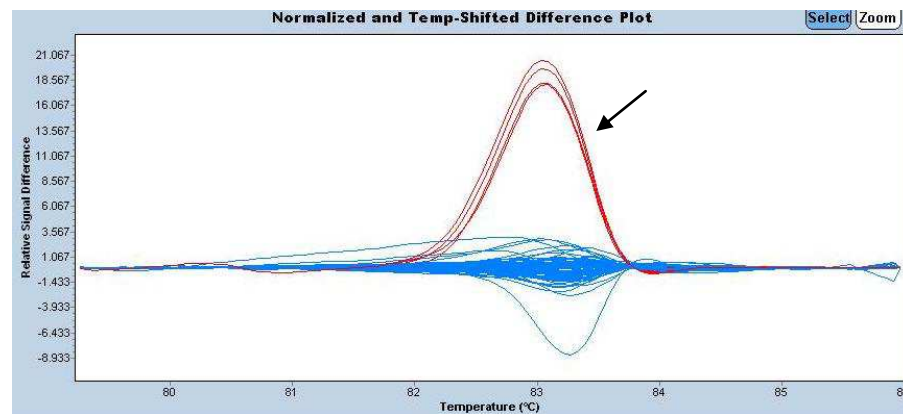
**Figure 16:**

**A.**

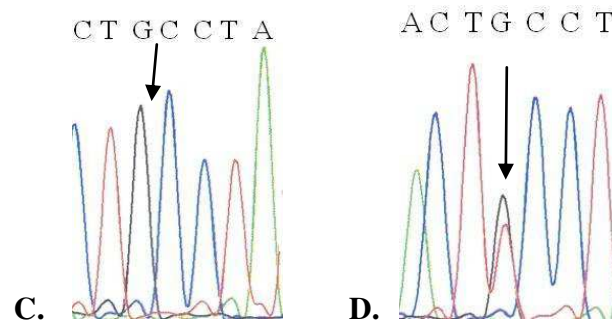


- A. “T<sub>m</sub> calling” analysis of the exon 17 of *MLH1* gene.** All peaks have approximately the same shape, thus there is only one PCR product in this analysis. Temperature is displayed on the x axis and the rate of fluorescence on the y axis.

**B.**



- B. Genetic analysis of the exon 17 of *MLH1* gene.** All curves have the same shape except of five curves representing samples bearing variants c. 1959 G>T. The relative signal of fluorescence represented by y axis is changing due to rising temperature and presence of mutations in *MLH1* gene.



**C.** The result of sequencing of exon 17 of *MLH1* gene without variants c. 1959 G>T (wild-type).

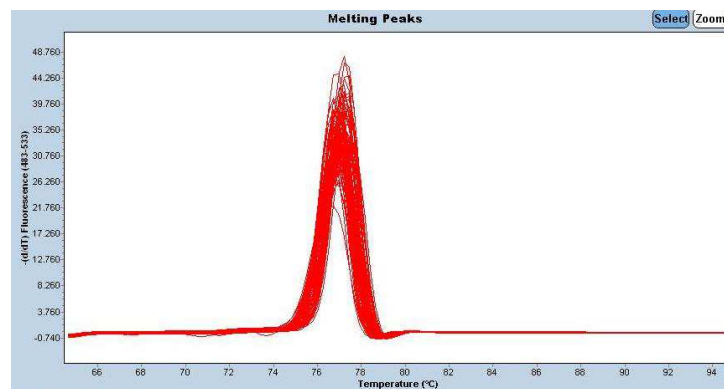
**D.** The result of sequencing of exon 17 of *MLH1* gene representing heterozygous carrier of gene variant c. 1959 G>T – two peaks at position c. 1959.

### D. Intronic variants in *MLH1* gene

We confirmed already known variant c. 1668-19 A>G in intron 14 in 73 DNA samples, 53 of them had heterozygous combination of alleles and the rest had homozygous combination of alleles (**Figure 17**).

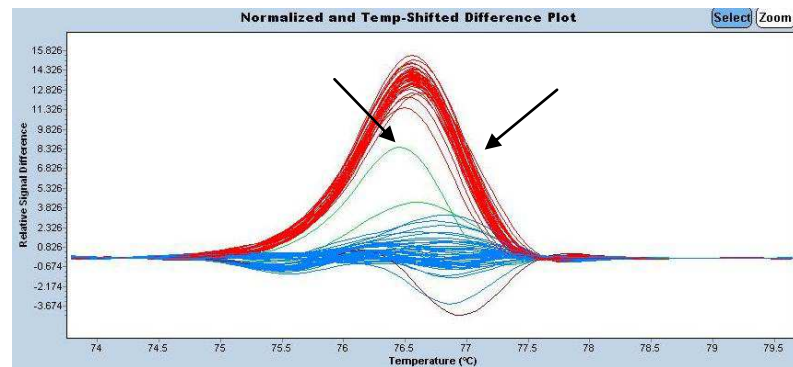
**Figure 17:**

**A.**

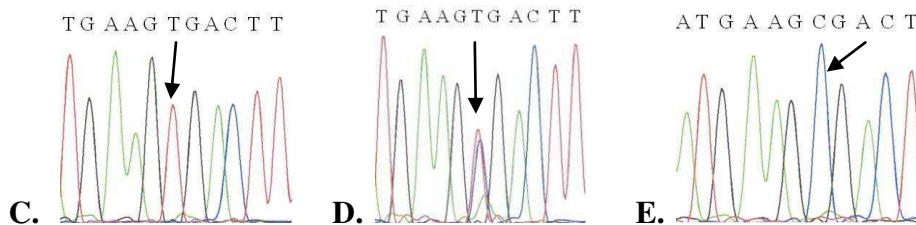


**A.** “T<sub>m</sub> calling” analysis of the intron 14 of *MLH1* gene. All peaks have approximately the same shape, thus there is only one PCR product in this analysis. Temperature is displayed on the x axis and the rate of fluorescence on the y axis.

**B.**



**B. Genetic analysis of the of the intron 14 *MLH1* gene.** All curves have the same shape except those 44 curves representing samples bearing variants c. 1668-19 A>G. The relative signal of fluorescence represented by y axis is changing due to rising temperature and presence of mutations in *MLH1* gene.



**C.** The result of sequencing of the intron 14 of *MLH1* gene without variants c. 1668-19 A>G (wild-type, REVERSE strand).

**D.** The result of sequencing of the intron. 14 of *MLH1* gene bearing heterozygous variants c. 1668-19 A>G – two peaks at position c. 204 (REVERSE strand).

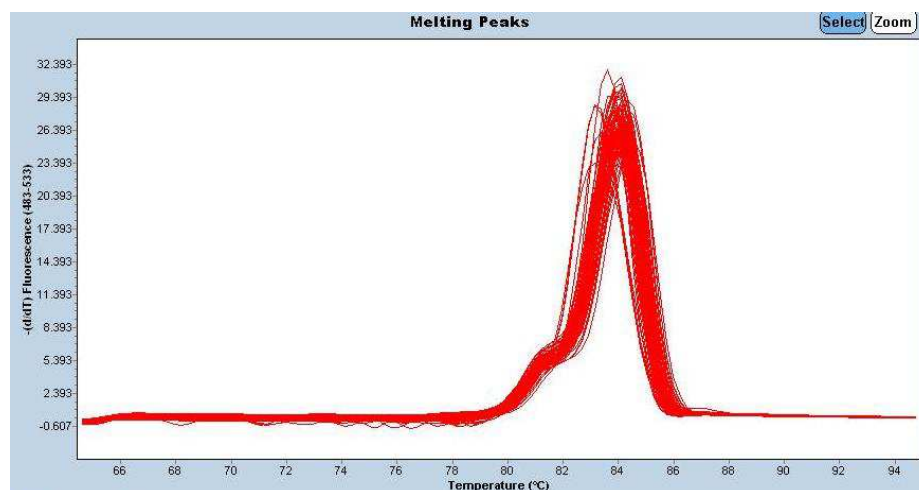
**E.** The result of sequencing of the intron 14 of *MLH1* gene bearing homozygous variants c. 1668-19 A>G (REVERSE strand).



Finally we found known polymorphism c. 1558+14 G>A in intron 13 in 6 DNA samples as heterozygous combination of alleles (**Figure 18**).

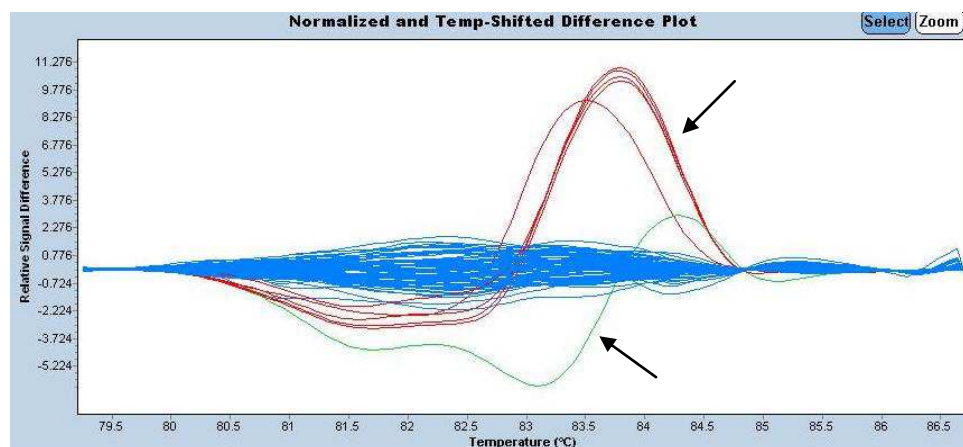
**Figure 18:**

**A.**

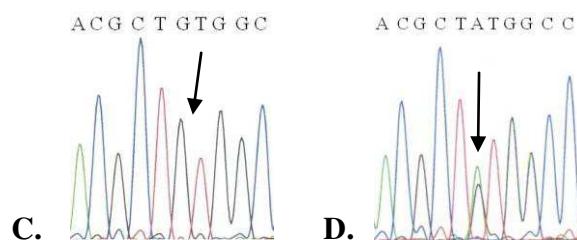


**A. “T<sub>m</sub> calling” analysis of the intron 13 of *MLH1* gene.** All peaks have approximately the same shape, thus there is only one PCR product in this analysis. Temperature is displayed on the x axis and the rate of fluorescence on the y axis.

**B.**



**B. Genetic analysis of the of the intron 13 *MLH1* gene.** All curves have the same shape except those six curves representing samples bearing variants c. 1558+14 G>A. The relative signal of fluorescence represented by y axis is changing due to rising temperature and presence of mutations in *MLH1* gene.



**Table 11: Summary of variants found in the *MLH1* gene.**

Variant in <i>MLH1</i>	Position	AA change	Genotypes	No. of cases	Type	Reference
<i>Somatic variants</i>						
c. 973 C>T	Exon 11	Arg325Trp	C/T	1	UV	-
<i>Germ-line variants</i>						
c. 1668-19 A>G	Intron 14	-	A/A	23	Polymorphism	Tannergård P. et al., 1995
			A/G	53		
			G/G	20		
c. 204 C>G	Exon 2	Ile68Met	C/G	1	UV	-
c. 655 A>G	Exon 8	Ile219Val	A/A	46	Polymorphism	Liu B. et al., 1995
			A/G	34		
			G/G	16		
c. 1733 A>G	Exon 16	Glu578Gly	A/G	1	Germ-line mutation	Tannergård P. et al., 1995
c. 1959 G>T	Exon 17	Leu653Leu	G/T	5	Polymorphism	Liu T. et al., 1998
c. 2146 G>A	Exon 19	Val716Met	G/A	2	Polymorphism	Genuardi et al., 1998
c.1558+14 G>A	Intron 13	-	G/A	6	Polymorphism	Ghimenti C. et al., 1999

UV = unclassified variant

MSI status was determined in all DNA samples, 9 of them were considered as MSI-H. Only one sample (141) bore somatic variant c. 973 C>T in exon 11. Other samples bore at least one SNP found in *MLH1* in this study (**Table 12**). Two DNA samples (193, 195) determined as MSI-H were obtained from one CRC patient.

Variant	c. 655 A>G, p. Ile219Val (ex 8)	c. 973 C>T, p. Arg325Trp (ex 11)	c. 1558+14 G>A (int 13)	c. 1668-19 A>G (int 14)	c. 1959 G>T, p. Leu653Leu (ex 17)
No. of sample					
29			•	•	
35	•			•	
119	•			•	
125	•			•	
141		•			
169	•			•	
173				•	
193				•	•
195				•	•

**Table 12: Gene variants found in DNA samples displaying MSI-H status.** The yellow cell represents somatic variant in *MLH1*, red cells represent two DNA samples obtained from one patient. Sample 193 - caecum, TNM III; BMI 23,12; 72 year old at the time of diagnosis; sample 195 - colon ascendens; TNM III.

## 6. Discussion

We have investigated mutations in *MLH1* gene in 96 patients with sporadic form of CRC and conferred them with the determined MSI status in each sample. We found two novel variants (c. 204 C>G, p. Ile68Met in exon 2; c. 973 C>T, p. Arg325Trp in exon 11), confirmed five known polymorphisms (c. 655 A>G, p. Ile219Val in exon 8; c. 1558+14 G>A in intron 13; c. 1668-19 A>G in intron 14; c. 1959 G>T, p. Leu653Leu in exon 17 and c. 2146 G>A, p. Val716Met in exon 19) and one mutation (c. 1733 A>G, p. Glu578Gly in exon 16), described previously in HNPCC (Tannergård et al., 1995). Five variants in exons 2, 8, 11, 16 and 19 are causing the change of amino acid, one variant (exon 17) do not change amino acid and two variants were found in intronic regions and were previously assigned changes that do not affect the splicing of *MLH1* (Tannergård P. et al., 1995; Ghimenti C. et al., 1999). MSI-High (MSI-H) status was determined in 9 samples from 99. It represents ~9 % of all samples; although according to literature 15 % of all sporadic CRCs display MSI-H status (Morán et al., 2010). Our result is slightly below that figure, probably due to the number of investigated subjects. Each sample characterized as MSI-H displayed at least one SNP in *MLH1* gene, except for sample 141 bearing only somatic variant c. 973 C>T in exon 11 (Sample 29= c. 1558+14 G>A, c. 1668-19 A>G; Samples 35, 119, 125, 169= c. 655 A>G, c. 1668-19 A>G; Sample 141= c. 973 C>T; Sample 173= c. 1668-19 A>G; Samples 193, 195= c. 1668-19 A>G, c. 1959 G>T). It is questionable if MSI in these patients is a consequence of changes in *MLH1* gene or in other MMR genes. An ongoing study on promoter methylation of MMR genes may elucidate this question. Only variant c. 1733 A>G, p. Glu578Gly in exon 16, found in one sample, has been attributed to be possibly pathogenic (Liu et al., 1999). However, the literature data regarding the above variant remain controversial (Guerrette et al., 1999; Lagerstedt-Robinson et al., 2007). Other known variants found in our study were many times described as neutral or not affecting substantially the function of MLH1 protein (Liu et al., 1995; Tannergård et al., 1995; Liu et al., 1998; Genuardi et al., 1998). Recently it has been revealed that polymorphisms could modulate the expression of *MLH1* gene. Mrkonjic et al. discovered that polymorphism MLH1-93 G>A is associated with

MSI-H colorectal cancers (Mrkonjic et al., 2010). They have also proved that the MSH2-118 T>C polymorphism is associated with clinical family history of CRC in Ontario patients (Mrkonjic et al., 2007). Further characterization of promoter polymorphisms and the cumulative effects that such alterations may have on gene expression warrants further studies. The above experiments along with miRNA regulations of MMR gene expressions may ultimately lead to the better understanding of the contribution of low-penetrant alleles to cancer incidence and progression and the overall CRC risk.

We have found two new variants in *MLH1* gene: Variant c. 204 C>G, p. Ile68Met in exon 2 (tumour was determined as MSS) causes the change of isoleucine to methionine. These amino acids are very similar in their width, thus it seems that this change do not necessarily affect the function of MLH1. This variant was also found in genomic DNA so it may not be ruled out that it confers with mutation detected in HNPCC cases (Papadopoulos et al., 1994). However, other investigation is warranted to prove this observation, mainly by family anamnesis, using Bethesda criteria for HNPCC (Piñol et al., 2005). Against the diagnosis of HNPCC speaks the age of the patient (61), since HNPCC is typical for younger patients (Oliviera et al., 2004). This variant may rather represent rare polymorphism. The second new variant (c. 973 C>T, p. Arg325Trp) was detected in the middle of exon 11 of one patient with MSI-H and it is putatively responsible for changing arginine to tryptophan. Exon 11 is encoding the ending region of MutS interaction domain of MLH1 thus it could affect its interaction ability with MSH2 or MSH6 (essential for function of MMR). Arginine has positive charge and is located on the inner surface of MLH1 while tryptophan is uncharged amino acid. Additional investigation of this change may provide answer to question, whether this change can cause malfunction of MLH1 on protein level. It is interesting that this patient has been two times diagnosed as CRC patient, first at the age of 53 and then there was a recurrence at the age of 80. The second tumour is presumably a relapse of the first one.

Another interesting casuistic is represented by one female patient of the age 73, who had four CRC tumours at the same time. Two were located in proximal colon and the other two in the distal one, from flexura lienalis to sigmoideum. All

DNA samples had c. 1959 G>T, p. Leu653Leu polymorphism in exon 17. Two samples were analyzed as MSI-H (with location in proximal colon); the other two samples were MSS. However, the above polymorphism does not affect the function of MLH1. The genetics of the four aforementioned tumours needs to be investigated. This patient is a unique example showing that CRC can arise at more locations in the colon or rectum and the developed tumours could exhibit both MSI and CIN phenotypes. Nowadays it is anticipated that CIN and MSI pathways are connected together. Moreover, mechanisms leading to MSI are not clear thus it is questionable if this patient had tumour displaying MSI as first tumour leading to development of the second one or vice versa.

We have found c. 655 A>G, p. Ile219Val polymorphism (rs1799977) in exon 8 in 50 DNA samples, which was previously described as gene variant with neutral effect (Trojan et al., 2002). Recently Picelli et al. proved that this variant is associated with an increased risk of colon and rectum cancer in Swedish population (Picelli et al., 2010). The investigation of the role of rs1799977 polymorphism on the CRC onset in the Czech population has very recently been carried out and the results became a part of an international meta-analysis (Picelli et al., submitted). However, no association of c. 655 A>G, p. Ile219Val polymorphism (rs1799977) with CRC risk has been discovered on almost 1000 patients of the Czech origin with sporadic CRC. Perhaps the high incidence of CRC in the Czech Republic is caused by random accumulation of polymorphisms leading to increased risk of CRC, their interactions with other gene variants (gene-gene interactions) or environmental factors. It is clear that high mortality is predominantly due to late diagnosis, but now it decreases mainly due to the improved health services and changes in the nutrition and life style.

CRC is a complex disease caused by genetic, epigenetic, life-style, environmental and other unknown processes. Recent investigations are focused mainly on other processes, such as SNPs, methylations or miRNA regulation. For instance, the SNPs near genes involved in DNA repair have been under-represented in the associations reported from genome-wide association studies (GWAS) of colorectal cancer susceptibility. Tomlinson et al. investigated the role of SNPs in 157 DNA repair genes in three CRC GWAS. Although no individual SNP showed

evidence of association, the set of SNPs was associated with the CRC risk. Mainly the *MLH1* promoter SNP -93G>A (rs1800734) and rare variants in *CHEK2* (1157T, del1100C) were associated with CRC risk (Tomlinson et al., 2012). The absence of disease-associated DNA repair SNPs in GWAS may be caused by more factors: many loci with individually very small effects on risk, rare alleles of moderate effect, MSI CRCs associated with specific variants. Perhaps the most relevant reason is that any alleles with more than very small functional effects are strongly selected and are very unlikely to drift up to polymorphic level (Tomlinson et al., 2012). GWAS in CRC have not discovered variation in pathways that could be described as strong candidates. In other cases, unexpected associations have been found. For example, transcription elongation factor EIF3H, ubiquitously expressed in all cells, has specific effects on CRC risk. The reasons for this are entirely unclear (Pittman et al., 2010).

It is also disputable, which processes contribute to development of defective DNA MMR. Diouf et al. proved that acute lymphoblastic leukaemia cells with low or undetectable MSH2 protein levels contained somatic mutations in 1-4 genes that regulate MSH2 degradation (Diouf et al., 2012). Prior investigations revealed that MSH2 had absent expression in ~15 % of sporadic CRC. For majority of these cases, the mechanism leading to low MSH2 protein was not identified (Kuismanen et al., 1999). According to Diouf et al., these cases may be caused by deletions of genes regulating MSH2 stability (Diouf et al., 2012).

Many authors also disputed if MSI-L CRCs represent distinct category of sporadic CRCs due to very similar clinico-pathological features with MSS CRCs. Recently, it has been reported that *MSH3* gene may contribute to development of MSI-L CRCs. The high frequency of LOH on chromosome 5q and the aberrant expression of MSH3 were described in some CRC tumours (Plaschke et al., 2012).

Nowadays, it is accepted that sporadic CRCs develop through two main pathways: CIN and MSI pathway. It is unclear, if these pathways are connected together and if there exists any cross-talk between them. Recent investigations revealed that the *in vitro* mutation frequency of *APC* gene is influenced by the MMR and base excision repair (BER) pathways. Cell lines Ibl-1261 and HCT116 with

defective DNA MMR activity have a lower mutation frequency in the *APC*. However, in the BER defective cell lines VACO425 and 2630, mutation frequencies in *APC* gene are higher (Turnbull et al., 2012).

CRC research has focused mainly on epigenetic mechanisms in the last decade. Mainly DNA methylation and miRNA regulation may significantly contribute to development of sporadic CRCs. Alterations in DNA methylation (DNAm) in cancer have been known for 25 years, including hypomethylation of oncogenes and hypermethylation of tumour suppressor genes (Feinberg and Tycko, 2004). Most studies on cancer methylation have assumed that functionally important DNAm will occur in promoters, and that most DNAm changes in cancer occur in CpG islands (Baylin and Ohm, 2006). However, Irizzary et al. showed that that most methylation alterations in colon cancer do not occur either in promoters, or in CpG islands, but in sequences up to 2 kb distant which we term “CpG island shores” (Irizzary et al., 2009). The connection between SNPs and proper function of miRNAs was recently documented. The presence of SNPs within the 30-untranslated regions of genes could affect the binding between miRNA and its target. Landi et al. investigated 12 SNPs that were genotyped in a case-control association study on 717 colorectal cases and 1171 controls from the Czech Republic. Statistically significant associations were found between the risk of CRC and the variant alleles of *KIAA0182* (rs709805) and *NUP210* genes (rs354476) (Landi et al., 2012). Also other groups showed that expression of particular miRNAs was decreased in sporadic CRCs, consequently contributing to its pathogenesis (Almeida et al., 2012; Faltejskova et al., 2012; Kuo et al., 2012; Zhang et al., 2012). In conclusion, it seems that CRC develops through many different mechanisms. Better understanding of molecular processes leading to CRC may contribute to the concept of individualized therapy.

This study is a part of more complex project on *MLH1* gene, which shall include also methylation analysis of the *MLH1* promoter and investigation of expression patterns of *MLH1*. Results from this project may provide more complex insight into molecular mechanisms leading to onset and progression of sporadic form of CRC. According to our results it seems that somatic mutations play minor role in



the development of sporadic form of CRC, at least within the group of patients investigated by us. It may be assumed that epigenetic mechanisms, mainly promoter methylation and miRNA regulation of MMR, may also play an important role.

## 7. Conclusions

Colorectal cancer represents one of the most frequent malignancies in the Czech population. In this study we analyzed somatic mutational events in *MLH1*, one of the MMR genes, in 99 samples obtained from 96 patients with the sporadic form of CRC.

- We successfully analyzed mutations in *MLH1* gene in 99 samples using HRM technique; suspicious samples were verified by automated sequencing.
- We found two new variants in *MLH1* gene (germ-line c. 204 C>G, p. Ile68Met in exon 2 and somatic one c. 973 C>T, p. Arg325Trp in exon 11) in relation to the sporadic CRC.
- MSI status was determined in all samples; MSI-H status was detected in 9 samples from 99. Only one sample bore somatic mutation in *MLH1* gene, other samples displaying MSI-H status bore at least one SNP in *MLH1*.

This study represents one of the first genetic analyses of somatic events in the *MLH1* gene on the Czech population. According to literature, *MLH1* promoter hypermethylation could be more important in sporadic CRC pathogenesis than somatic mutations. Further investigation for better understanding of molecular processes leading to sporadic form of CRC displaying MSI is warranted.

## 8. References

- Aebi S., Fink D., Gordon R., Kim H.K., Zheng H., Fink J.L., Howell S.B.** (1997) Resistance to cytotoxic drugs in DNA mismatch repair-deficient cells. *Clin Cancer Res*;3: 1763–1767.
- Allen D.J., Makhov A., Grilley M., Taylor J., Thresher R., Modrich P., Griffith J.D.** (1997) MutS mediates heteroduplex loop formation by a translocation mechanism. *EMBO J.* 16, 4467–4476.
- Almeida M.I., Nicoloso M.S., Zeng L., Ivan C., Spizzo R., Gafà R., Xiao L., Zhang X., Vannini I., Fanini F., Fabbri M., Lanza G., Reis R.M., Zweidler-McKay P.A., Calin G.A.** (2012) Strand-Specific miR-28-5p and miR-28-3p Have Distinct Effects in Colorectal Cancer Cells. *Gastroenterology*. 142(4):886-896.e9. Epub 2012 Jan 10.
- Al-Tassan N., Chmiel N.H., Maynard J., Fleming N., Livingston A.L. Williams G.T., Hodges A.K., Davies D.R., David S.S., Sampson J.R., Cheadle J.P.** (2002) Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. *Nat Genet* 30: 227–232.
- Ban C., Junop M., Yang W.** (1999) Transformation of MutL by ATP binding and hydrolysis: a switch in DNA mismatch repair. *Cell* 97, 85–97.
- Banerjea A., Ahmed S., Hands R.E., Huang F., Han X., Shaw P.M., Feakins R., Bustin S.A., Dorudi S.** (2004) Colorectal cancers with microsatellite instability display mRNA expression signatures characteristic of increased immunogenicity. *Mol. Cancer*, 3, 21.
- Barker N.** (2008) The canonical Wnt/beta-catenin signalling pathway. *Methods Mol Biol*; 468: 5-15.
- Baylin S.B., Ohm J.E.** (2006) Epigenetic gene silencing in cancer - a mechanism for early oncogenic pathway addiction? *Nat Rev Cancer* 6:107–16.
- Beggs A.D., Latchford A.R., Vasen H.F.A.** (2010) Peutz-Jeghers syndrome: a systematic review and recommendations for management. *Gut*, vol. 59, no. 7, pp. 975–986.
- Benatti P., Gafà R., Barana D., Marino M., Scarselli A., Pedroni M., Maestri I., Guerzoni L., Roncucci L., Menigatti M., Roncari B., Maffei S., Rossi G., Ponti G., Santini A., Losi L., Di Gregorio C., Olini C., Ponz de Leon M., Lanza G.** (2005) Microsatellite instability and colorectal cancer prognosis. *Clin Cancer Res*; 11: 8332-8340.
- Benchabane H., Ahmed Y.** (2009) The adenomatous polyposis coli tumor suppressor and Wnt signaling in the regulation of apoptosis. *Advances in Experimental Medicine and Biology*, vol. 656, pp. 75–84.
- Boland C.R., Thibodeau S.N., Hamilton S.R., Sidransky D., Eshleman J.R., Burt R.W., Meltzer S.J., Rodriguez-Bigas M.A., Fodde R., Ranzani G.N., Srivastava S.** (1998) A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res*. 58(22):5248-57.

- Boland C.R., Goel A.** (2010) Microsatellite instability in colorectal cancer. *Gastroenterology*. 138(6):2073-2087.e3.
- Buhard O., Cattaneo F., Wong Y.F., Yim S.F., Friedman E., Flejou J.F., Duval A., Hamelin R.** (2006) Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumors. *J Clin Oncol*. 24(2):241-51.
- Campa D., Pardini B., Naccarati A., Vodickova L., Novotny J., Försti A., Hemminki K., Barale R., Vodicka P., Canzian F.** (2008) A gene-wide investigation on polymorphisms in the ABCG2/BRCP transporter and susceptibility to colorectal cancer. *Mutat Res*. 645(1-2):56-60.
- Campa D., Pardini B., Naccarati A., Vodickova L., Novotny J., Steinke V., Rahner N., Holinski-Feder E., Morak M., Schackert H.K., Görgens H., Kötting J., Betz B., Kloor M., Engel C., Büttner R., Propping P., Försti A., Hemminki K., Barale R., Vodicka P., Canzian F.** (2010) Polymorphisms of genes coding for ghrelin and its receptor in relation to colorectal cancer risk: a two-step gene-wide case-control study. *BMC Gastroenterol*. 10:112.
- Cannavo E., Marra G., Sabates-Bellver J., Menigatti M., Lipkin S.M., Fischer F., Cejka P., Jiricny J.** (2005) Expression of the MutL homologue hMLH3 in human cells and its role in DNA mismatch repair. *Cancer Res*. 65, 10759–10766.
- Carethers J.M. Smith E.J., Behling C.A., Nguyen L., Tajima A., Doctolero R.T., Cabrera B.L., Goel A., Arnold C.A., Miyai K., Boland C.R.** (2004) Use of 5-fluorouracil and survival in patients with microsatellite unstable colorectal cancer. *Gastroenterology*, 126, 394–401.
- Cheah P.Y., Wong YU.H., Chau Y.P.** (2009) Germline bone morphogenesis protein receptor 1A mutation causes colorectal tumorigenesis in hereditary mixed polyposis syndrome. *American Journal of Gastroenterology*, vol. 104, no. 12, pp. 3027–3033.
- Chen Y.Q., Hsieh J.T., Yao F., Fang B., Pong R.C., Cipriano S.C., Krepulat F.** (1999) Induction of apoptosis and G2/M cell cycle arrest by DCC. *Oncogene*; 18: 2747-2754.
- Cunningham D., Atkin W., Lenz H.J., Lynch H.T., Minsky B., Nordlinger B., Starling N.** (2010) Colorectal cancer. *Lancet* 375 (9719): 1030–47.
- Diouf B., Cheng Q., Krynetskaia N.F., Yang W., Cheok M., Pei D., Fan Y., Cheng C., Krynetskiy E.Y., Geng H., Chen S., Thierfelder W.E., Mullighan C.G., Downing J.R., Hsieh P., Pui C.H., Relling M.V., Evans W.E.** (2011) Somatic deletions of genes regulating MSH2 protein stability cause DNA mismatch repair deficiency and drug resistance in human leukemia cells. *Nat Med*. 17(10):1298-303. doi: 10.1038/nm.2430.
- Dusek L., Muzik J., Kubasek M., Koptikova J., Zaloudik J., Vyzula R.** (2005) Epidemiologie zhoubných nádorů v České republice. Masarykova univerzita.
- Eshleman J.R., Casey G., Kochera M.E., Sedwick W.D., Swinler S.E., Veigl M.L., Willson J.K., Schwartz S., Markowitz S.D.** (1998) Chromosome number and structure both are markedly stable in RER colorectal cancers and are not destabilized by mutation of p53. *Oncogene*; 17: 719-725.
- Esteve P.O., Chin H.G., Pradhan S.** (2005) Human maintenance DNA (cytosine-5)-methyltransferase and p53 modulate expression of p53-repressed promoters. *Proc Natl Acad Sci U S A*. Jan 25;102(4):1000-5. Epub 2005 Jan 18.
- Fallik D., Borrini F., Boige V., Viguier J., Jacob S., Miquel C., Sabourin J.C., Ducreux M., Praz F.** (2003) Microsatellite instability is a predictive factor of the tumor response to irinotecan in patients with advanced colorectal cancer. *Cancer Res*;63:5738–5744.

- Faltejskova P., Svoboda M., Srutova K., Mlcochova J., Besse A., Nekvindova J., Radova L., Fabian P., Slaba K., Kiss I., Vyzula R., Slaby O.** (2012) Identification and functional screening of microRNAs highly deregulated in colorectal cancer. *J Cell Mol Med.* doi: 10.1111/j.1582-4934.2012.01579.x. [Epub ahead of print]
- Fearon E.R., Vogelstein B.** (1990) A genetic model for colorectal tumorigenesis. *Cell* 61 759–767.
- Feinberg A.P., Tycko B.** (2004) The history of cancer epigenetics. *Nat Rev Cancer.* 4:143–53.
- Ferlay J., Autier P., Boniol M., Heanue M., Colombet M.** (2007) Estimates of the cancer incidence and mortality in Europe in 2006. *Ann Oncol* 18: 581–592.
- Frattoni M.B.D., Suardi S., Oggionni M., Alberici P., Radice P., Costa A. Daidone M.G., Leo E., Pilotti S., Bertario L., Pierotti M.A.** (2004) Different genetic features associated with colon and rectal carcinogenesis. *Clin Cancer Res*;10: 4015e21.
- Fuentealba L.C., Eivers E., Ikeda A., Hurtado C., Kuroda H., Pera E.M., De Robertis E.M.** (2007) Integrating patterning signals: Wnt/GSK3 regulates the duration of the BMP/Smad1 signal. *Cell*; 131:980-993.
- Genschel J., Littman S.J., Drummond J.T., Modrich P.** (1998) Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. *J. Biol. Chem.* 273, 19895–19901.
- Genuardi M., Anti M., Capozzi E., Leonardi F., Fornasarig M., Novella E., Bellacosa A., Valenti A., Gasbarrini G.B., Roncucci L., Benatti P., Percesepe A., Ponz de Leòn M., Coco C., de Paoli A., Valentini M., Boiocchi M., Neri G., Viel A.** (1998) MLH1 and MSH2 constitutional mutations in colorectal cancer families not meeting the standard criteria for hereditary nonpolyposis colorectal cancer. *Int J Cancer.* 75(6):835-9.
- Ghimenti C., Tannergård P., Wahlberg S., Liu T., Giulianotti P.G., Mosca F., Fornaciari G., Bevilacqua G., Lindblom A., Caligo M.A.** (1999) Microsatellite instability and mismatch repair gene inactivation in sporadic pancreatic and colon tumours. *Br J Cancer.* 80(1-2):11-6.
- Greenson J.K., Bonner J.D., Ben-Yzhak O., Cohen H.I., Miselevich I., Resnick M.B., Trougouboff P., Tomsho L.D., Kim E., Low M., Almog R., Rennert G., Gruber S.B.** (2003) Phenotype of microsatellite unstable colorectal carcinomas: Well-differentiated and focally mucinous tumors and the absence of dirty necrosis correlate with microsatellite instability. *Am J Surg Pathol*; 27: 563-570.
- Gryfe R., Kim H., Hsieh E.T., Aronson M.D., Holowaty E.J., Bull S.B., Redston M., Gallinger S.** (2000) Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. *N Engl J Med*; 342: 69-77.
- Guerrette S., Acharya S., Fishel R.** (1999) The interaction of the human MutL homologues in hereditary nonpolyposis colon cancer. *J Biol Chem.* 274(10):6336-41.
- Guidoboni M., Gafà R., Viel A., Doglioni C., Russo A., Santini A., Del Tin L., Macri E., Lanza G., Boiocchi M., Dolcetti R.** (2001) Microsatellite instability and high content of activated cytotoxic lymphocytes identify colon cancer patients with a favorable prognosis. *Am J Pathol.* 159(1):297-304.
- Haigis K.M., Kendall K.R., Wang Y., Cheung A., Haigis M.C., Glickman J.N., Niwa-Kawakita M., Sweet-Cordero A., Sebolt-Leopold J., Shannon K.M., Settleman J., Giovannini M., Jacks T.** (2008) Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon. *Nat Genet*; 40: 600-608.

- Half E., Bercovich D., Rozen P.** (2009) Familial adenomatous polyposis. *Orphanet Journal of Rare Diseases*, vol. 4, no. 1, article 22.
- Hoeijmakers J.H.** (2001) Genome maintenance mechanisms for preventing cancer. *Nature*; 411: 366-734.
- Howe J.R., Roth S., Ringold J.C.** (1998) Mutations in the SMAD4/DPC4 gene in juvenile polyposis. *Science*, vol. 280, no. 5366, pp. 1086–1088.
- Howe J.R., Bair J.L., Sayed M.G.** (2001) Germline mutations of the gene encoding bone morphogenetic protein receptor 1A in juvenile polyposis. *Nature Genetics*, vol. 28, no. 2, pp. 184–187.
- Imai K., Yamamoto H.** (2008) Carcinogenesis and microsatellite instability: the interrelationship between genetics and epigenetics. *Carcinogenesis*; 29: 673-680.
- Ionov Y., Peinado M.A., Malkhosyan S., Shibata D., Perucho M.** (1993) Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363: 558–561.
- Irizarry R.A., Ladd-Acosta C., Wen B., Wu Z., Montano C., Onyango P., Cui H., Gabo K., Rongione M., Webster M., Ji H., Potash J.B., Sabunciyan S., Feinberg A.P.** (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet.* 41(2):178-86.
- Issa J.P.** (2004) CpG island methylator phenotype in cancer. *Nat. Rev. Cancer*, 4, 988–993.
- Jaspersion K.W., Tuohy T.M., Neklason D.W., Burt R.W.** (2010) Hereditary and familial colon cancer. *Gastroenterology*, vol. 138, no. 6, pp. 2044–2058.
- Jass J.R.** (2002) Pathogenesis of colorectal cancer. *Surg Clin North Am*; 82: 891-904.
- Jass J.R., Walsh M.D., Barker M., Simms L.A., Young J., Leggett B.A.** (2002) Distinction between familial and sporadic forms of colorectal cancer showing DNA microsatellite instability. *Eur J Cancer*; 38: 858-866.
- Jemal A., Bray F., Center M.M., Ferlay J., Ward E., Forman D.** (2011) Global cancer statistics. *CA Cancer J Clin* 61 (2): 69–90.
- Jenne D.E., Reimann H., Nezu J.I.** (1998) Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. *Nature Genetics*, vol. 18, no. 1, pp. 38–43.
- Jiricny J.** (2006) The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol.*; 7(5):335-46.
- Jo W.S., Carethers J.M.** (2006) Chemotherapeutic implications in microsatellite unstable colorectal cancer. *Cancer Biomark*; 2:51e60.
- Kadyrov F.A., Dzantiev L., Constantin N., Modrich P.** (2006) Endonucleolytic function of MutL alpha in human mismatch repair. *Cell* 126, 297–308.
- Knudson A.** (1971) Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A.* Apr;68(4):820-3.
- Kondo E., Horii A., Fukushima S.** (2001) The interacting domains of three MutL heterodimers in man: hMLH1 interacts with 36 homologous amino acid residues within hMLH3, hPMS1 and hPMS2. *Nucleic Acids Res.* 29, 1695–1702.

- Kopacova M., Tacheci I., Rejchrt S., Bures J.** (2009) Peutz-Jeghers syndrome: diagnostic and therapeutic approach. *World Journal of Gastroenterology*, vol. 15, no. 43, pp. 5397–5408.
- Kranenburg O.** (2005) The KRAS oncogene: past, present and future. *1756(2)*:81-2. Epub 2005 Oct 25.
- Kuismanen S.A., Holmberg M.T., Salovaara R., Schweizer P., Aaltonen L.A., de La Chapelle A., Nyström-Lahti M., Peltomäki P.** (1999) Epigenetic phenotypes distinguish microsatellite-stable and -unstable colorectal cancers. *Proc. Natl. Acad. Sci. USA*. 1999; 96:12661–12666.
- Kuismanen S.A., Holmberg M.T., Salovaara R., de la Chapelle A., Peltomäki P.** (2000) Genetic and epigenetic modification of MLH1 accounts for a major share of microsatellite-unstable colorectal cancers. *Am J Pathol*. May;156(5):1773-9.
- Kuo T.Y., Hsi E., Yang I.P., Tsai P.C., Wang J.Y., Juo S.H.** (2012) Computational analysis of mRNA expression profiles identifies microRNA-29a/c as predictor of colorectal cancer early recurrence. *PLoS One*. 7(2):e31587. Epub 2012 Feb 13.
- Labianca R., Beretta G.D., Kildani B.** (2010) Colon cancer. *Critical Reviews in Oncology/Hematology*, vol. 74, no. 2, pp. 106–133.
- Lagerstedt-Robinson K., Liu T., Vandrovcova J., Halvarsson B., Clendenning M., Frebourg T., Papadopoulos N., Kinzler K.W., Vogelstein B., Peltomäki P., Kolodner R.D., Nilbert M., Lindblom A.** (2007) Lynch syndrome (hereditary nonpolyposis colorectal cancer) diagnostics. *J Natl Cancer Inst*. 99(4):291-9.
- Lahue R.S., Au K.G., Modrich P.** (1989) DNA mismatch correction in a defined system. *Science* 245, 160–164.
- Laiho P., Launonen V., Lahermo P., Esteller M., Guo M., Herman J.G., Mecklin J.P., Järvinen H., Sistonen P., Kim K.M., Shibata D., Houlston R.S., Aaltonen L.A.** (2002) Low-level microsatellite instability in most colorectal carcinomas. *Cancer Res*; 62: 1166-1170.
- Landi D., Gemignani F., Pardini B., Naccarati A., Garritano S., Vodicka P., Vodickova L., Canzian F., Novotny J., Barale R., Landi S.** (2012) Identification of candidate genes carrying polymorphisms associated with the risk of colorectal cancer by analyzing the colorectal mutome and microRNAome. *Cancer*. doi: 10.1002/cncr.27435. [Epub ahead of print]
- Laurent-Puig P., Agostini J., Maley K.** (2010) Colorectal oncogenesis. *Bull Cancer*. 97(11):1311-21.
- Lengauer C., Kinzler K.W., Vogelstein B.** (1998) Genetic instabilities in human cancers. *Nature*. 396(6712):643-9.
- Leslie A., Carey F.A., Pratt N.R., Steele R.J.** (2002) The colorectal adenoma-carcinoma sequence. *Br J Surg*; 89: 845-860.
- Liaw D., Marsh D. J., Li J.** (1997) Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nature Genetics*, vol. 16, no. 1, pp. 64–67.

- Liu B., Nicolaides N.C., Markowitz S., Willson J.K., Parsons R.E., Jen J., Papadopoulos N., Peltomäki P., de la Chapelle A., Hamilton S.R. Kinzler K.W., Vogelstein B.** (1995) Mismatch repair gene defects in sporadic colorectal cancers with microsatellite instability. *Nat Genet.* 9(1):48-55.
- Liu T., Wahlberg S., Rubio C., Holmberg E., Grönberg H., Lindblom A.** (1998) DGGE screening of mutations in mismatch repair genes (hMSH2 and hMLH1) in 34 Swedish families with colorectal cancer. *Clin Genet.* 53(2):131-5.
- Livesey F.J., Hunt S.P.** (1997) Netrin and netrin receptor expression in the embryonic mammalian nervous system suggests roles in retinal, striatal, nigral, and cerebellar development. *Molecular and Cellular Neuroscience* 8, 417–429.
- Lynch H.T., Lynch P. M., Lanspa S. J., Snyder C. L., Lynch J. F., Boland C. R.** (2009) Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clinical Genetics*, vol. 76, no. 1, pp. 1–18.
- MacKay S.L.D., Yaswen L.R., Tarnuzzer R.W., Moldawer L.L., Bland K.I., Copeland E.M., Schultz G.S.** (1995) Colon cancer cells that are not growth inhibited by TGF- $\beta$  lack functional type I and type II TGF- $\beta$  receptors. *Ann Surg*; 221:767-77.
- Markowitz S., Wang L., Myeroff L., Parsons R., Sun L., Lutterbaugh L., Fan R.S., Zborowska E., Kinzler K.W., Vogelstein B. et al.** (1995) Inactivation of the type II TGF- $\beta$  receptor in colon cancer cells with microsatellite instability. *Science*, 268:1336-8.
- Marra G., Jiricny J.** (2005) DNA mismatch repair and colon cancer. *Adv Exp Med Biol.* 570:85-123.
- Marsischky G.T., Filosi N., Kane M.F., Kolodner R.** (1996) Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. *Genes Dev.* 10, 407–420.
- Massague J., Seoane J., Wotton D.** (2005) Smad transcription factors. *Genes Dev.* 19 (23): 2783–810.
- Migliore L., Migheli F., Spisni R., Coppedè F.** (2011) Genetics, cytogenetics, and epigenetics of colorectal cancer. *J Biomed Biotechnol.* 2011:792362. Epub 2011 Feb 14.
- Modrich P., Lahue R.** (1996) Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu Rev Biochem.* 65:101-33.
- Moran A., Ortega P., de Juan C., Fernandez-Marcelo T., Frías C., Sanchez-Pernaute A., Torres A.J., Diaz-Rubio E., Iniesta P., Benito M.** (2010) Differential colorectal carcinogenesis: Molecular basis and clinical relevance. *World J Gastrointest Oncol.*; 2(3):151-8.
- Mori S., Ogata Y., Shirouzu K.** (2004) Biological features of sporadic colorectal carcinoma with high-frequency microsatellite instability: special reference to tumor proliferation and apoptosis. *Int J Clin Oncol*;9:322e9.
- Mrkonjic M., Raptis S., Green R. C., Monga N., Daftary D., Dicks E., Younghusband H.B., Parfrey P.S., Gallinger S.S., McLaughlin J.R., Knight J.A., Bapat B.** (2007) MSH2 118T>C and MSH6 159C>T promoter polymorphisms and the risk of colorectal cancer. *Carcinogenesis* 28: 2575–2580.



**Mrkonjic M., Roslin N.M., Greenwood C.M., Raptis S., Pollett A., Laird P.W., Pethe V.V., Chiang T., Daftary D., Dicks E., Thibodeau S.N., Gallinger S., Parfrey P.S., Younghusband H.B., Potter J.D., Hudson T.J., McLaughlin J.R., Green R.C., Zanke B.W., Newcomb P.A., Paterson A.D., Bapat B.** (2010) Specific variants in the MLH1 gene region may drive DNA methylation, loss of protein expression, and MSI-H colorectal cancer. *PLoS One*. Oct 13;5(10):e13314.

**Naccarati A., Polakova V., Pardini B., Vodickova L., Hemminki K., Kumar R., Vodicka P.** (2012) Mutations and polymorphisms in TP53 gene--an overview on the role in colorectal cancer. *Mutagenesis*. 27(2):211-8.

**Nevers P., Spatz H.C.** (1975) *Escherichia coli* mutants uvr D and uvr E deficient in gene conversion of lambda-heteroduplexes. *Mol Gen Genet*. 27;139(3):233-43.

**O'Brien V., Brown R.** (2006) Signalling cell cycle arrest and cell death through the MMR system. *Carcinogenesis* 27, 682–692.

**Ogino S., Kawasaki T., Kirkner G.J., Loda M., Fuchs C.S.** (2006) CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. *J. Mol. Diagn.*, 8, 582–588.

**Oliveira C., Westra J.L., Arango D., Ollikainen M., Domingo E., Ferreira A., Velho S., Niessen R., Lagerstedt K., Alhopuro P., Laiho P., Veiga L., Teixeira M.R., Ligtenberg M., Kleibeuker J.H., Sijmons R.H., Plukker J.T., Imai K., Lage P., Hamelin R., Albuquerque C., Schwartz S. Jr, Lindblom A., Peltomaki P., Yamamoto H., Aaltonen L.A., Seruca R., Hofstra R.M.** (2004) Distinct patterns of KRAS mutations in colorectal carcinomas according to germline mismatch repair defects and hMLH1 methylation status. *Hum Mol Genet*; 13: 2303-2311.

**Oren M., Rotter V.** (2010) Mutant p53 gain-of-function in cancer. *Cold Spring Harb Perspect Biol*. Feb;2(2):a001107.

**Papadopoulos N., Nicolaidis N., Wei Y., Ruben S., Carter K., Rosen C., Haseltine W., Fleischmann R., Fraser C., Adams M.** (1994) Mutation of a mutL homolog in hereditary colon cancer. *Science* 263 (5153): 1625–9.

**Pawlik T.M., Raut C.P., Rodriguez-Bigas M.A.** (2004) Colorectal carcinogenesis: MSI-H versus MSI-L. *Dis Markers*; 20: 199-206.

**Peltomaki, P.** (2003) Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J Clin Oncol*. Vol. 15;21(6):1174-9.

**Picelli S., Zajac P., Zhou X.L., Edler D., Lenander C., Dalén J., Hjern F., Lundqvist N., Lindfors U., Pählman L., Smedh K., Törnqvist A., Holm J., Janson M., Andersson M., Ekelund S., Olsson L., Lundeberg J., Lindblom A.** (2010) Common variants in human CRC genes as low-risk alleles. *European Journal of Cancer*. Vol. 46, Issue 6, p. 1041-1048.

**Picelli S., Bermejo J.S., Chang-Claude J., Hoffmeister M., Fernanández-Rozadilla C., Carracedo A., Castells A., Castellví-Bel S., The EPICOLON Consortium, Naccarati A., Pardini B., Vodickova L., Müller H., Talseth-Palmer B.A., Stibbard G., Peterlongo P., Nici C., Veneroni S., Daidone M.G., Li L., Tenesa A., van Wezel T., Farrington S.M., Tomlinson I., Moreno V., Wijnen J., Dunlop M., Radice P., Scott R.J., Vodicka P., Ruiz-Ponte C., Brenner H., Buch S., Lindblom A.** (submitted 2012) Shrinkage of association signals with increasing sample size illustrated by a meta-analysis of six polymorphisms and colorectal cancer susceptibility.

- Pinol V., Castells A., Andreu M., Castellví-Bel S., Alenda C., Llor X., Xicola R.M., Rodríguez-Moranta F., Payá A., Jover R., Bessa X.** (2005). Accuracy of revised Bethesda guidelines, microsatellite instability, and immunohistochemistry for the identification of patients with hereditary nonpolyposis colorectal cancer. *JAMA* 293 (16): 1986–1994.
- Pittman A.M., Naranjo S., Webb E., Broderick P., Lips E.H., Wezel T.van, Morreau H., Sullivan K., Fielding S., Twiss P., Vijayakrishnan J., Casares F., Qureshi M., Gómez-Skarmeta J.L., Houlston R.S.** (2009) The colorectal cancer risk at 18q21 is caused by a novel variant altering SMAD7 expression. *Genome Res.* 19: 987-993
- Pittman A.M., Naranjo S., Jalava S.E. Twiss P., Ma Y., Olver B., Lloyd A., Vijayakrishnan J., Qureshi M., Broderick P., van Wezel T., Morreau H., Tuupanen S., Aaltonen L.A., Alonso M.E., Manzanares M., Gavilán A., Visakorpi T., Gómez-Skarmeta J.L., Houlston R.S.** (2010) Allelic variation at the 8q23.3 colorectal cancer risk locus functions as a cis-acting regulator of EIF3H. *PLoS Genet.*, 6, e1001126.
- Plaschke J., Preußler M., Ziegler A., Schackert H.K.** (2012) Aberrant protein expression and frequent allelic loss of MSH3 in colorectal cancer with low-level microsatellite instability. *Int J Colorectal Dis.* 2012 Jan 10. [Epub ahead of print]
- Polakova V., Pardini B., Naccarati A., Landi S., Slyskova J., Novotny J., Vodickova L., Bermejo J.L., Hanova M., Smerhovsky Z., Tulupova E., Kumar R., Hemminki K., Vodicka P.** (2009) Genotype and haplotype analysis of cell cycle genes in sporadic colorectal cancer in the Czech Republic. *Hum. Mutat.*, 30, 661–668.
- Popat,S., Hubner R., Houlston R.S.** (2005) Systematic review of microsatellite instability and colorectal cancer prognosis. *J. Clin. Oncol.*, 23, 609–618.
- Rajagopalan H., Nowak M.A., Vogelstein B., Lengauer C.** (2003) The significance of unstable chromosomes in colorectal cancer. *Nature Reviews Cancer*, vol. 3, no. 9, pp. 695–701.
- Raut C.P., Pawlik T.M., Rodriguez-Bigas M.A.** (2004) Clinicopathologic features in colorectal cancer patients with microsatellite instability. *Mutat Res*; 568: 275-282.
- Ribic C.M., Sargent D.J., Moore M.J., Thibodeau S.N., French A.J., Goldberg R.M., Hamilton S.R., Laurent-Puig P., Gryfe R., Shepherd L.E., Tu D., Redston M., Gallinger S.** (2003) Tumor microsatellite instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N Engl J Med*;349:247–257.
- Rocha S., Martin A.M., Meek D.W., Perkins N.D.** (2003) P53 represses cyclin D1 transcription through down regulation of Bcl-3 and inducing increased association of the p52 NF-kappaB subunit with histone deacetylase 1. *Mol Cell Biol.* Jul;23(13):4713-27.
- Rouleau E., Lefol C., Bourdon V., Coulet F., Noguchi T., Soubrier F., Bieche I., Olschwang S., Sobol H., Lidereau R.** (2009) Quantitative PCR high-resolution melting (qPCR-HRM) curve analysis, a new approach to simultaneously screen point mutations and large rearrangements: application to MLH1 germline mutations in Lynch syndrome. *Hum Mutat* 30, 867–875.
- Rubio C. A., Stemme S., Jaramillo E., Lindblom A.** (2006) Hyperplastic polyposis coli syndrome and colorectal carcinoma. *Endoscopy*, vol. 38, no. 3, pp. 266–270.
- Rusan N.M., Peifer M.** (2008) Original CIN: reviewing roles for APC in chromosome instability. *J Cell Biol*; 181: 719-726.

- Russo A., Bazan V., Iacopetta B., Kerr D., Soussi T., Gebbia N.** (2005) The TP53 colorectal cancer international collaborative study on the prognostic and predictive significance of p53 mutation: influence of tumor site, type of mutation, and adjuvant treatment. *Journal of Clinical Oncology*, vol. 23, no. 30, pp. 7518–7528.
- Samowitz W.S., Curtin K., Ma K.N., Schaffer D., Coleman L.W., Leppert M., Slattery M.L.** (2001) Microsatellite instability in sporadic colon cancer is associated with an improved prognosis at the population level. *Cancer Epidemiol Biomarkers Prev*; 10: 917-923.
- Samowitz W.S., Albertsen H., Herrick J., Levin T.R., Sweeney C., Murtaugh M.A., Wolff R.K., Slattery M.L.** (2005) Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. *Gastroenterology*, 129, 837–845.
- Shekarabi M., Kennedy T.E.** (2002) The netrin-1 receptor DCC promotes filopodia formation and cell spreading by activating Cdc42 and Rac1. *Mol Cell Neurosci*; 19: 1-17.
- Smakman N., Borel Rinkes I.H., Voest E.E., Kranenburg O.** (2005) Control of colorectal metastasis formation by K-Ras. *Biochim Biophys Acta*; 1756: 103-114.
- Søreide K., Janssen E.A., Søiland H., Kørner H., Baak J.P.** (2006) Microsatellite instability in colorectal cancer. *Br J Surg*;93:395e406.
- Stojic L., Brun R., Jiricny J.** (2004) Mismatch repair and DNA damage signalling. *DNA Repair (Amst.)* 3, 1091–1101.
- Suraweera N., Duval A., Reperant M., Vaury C., Furlan D., Leroy K., Seruca R., Iacopetta B., Hamelin R.** (2002) Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology*, 123, 1804–1811.
- Takagi Y., Kohmura H., Futamura M., Kida H., Tanemura H., Shimaoka K., Saji S.** (1996) Somatic alterations of the DPC4 gene in human colorectal cancers in vivo. *Gastroenterology* 111, 1369-1372.
- Takayama T., Miyanishi K., Hayashi T., Sato Y., Niitsu Y.** (2006) Colorectal cancer: genetics of development and metastasis. *J Gastroenterol*; 41: 185-192.
- Tannergård P., Lipford J.R., Kolodner R., Frödin J.E., Nordenskjöld M., Lindblom A.** (1995) Mutation screening in the hMLH1 gene in Swedish hereditary nonpolyposis colon cancer families. *Cancer Res.* 55(24):6092-6.
- ten Dijke P., Hill C.S.** (2004) New insights into TGF-beta-Smad signalling. *Trends Biochem. Sci.* 29: 265–273.
- Tomlinson I.P., Houlston R.S., Montgomery G.W., Sieber O.M., Dunlop M.G.** (2012) Investigation of the effects of DNA repair gene polymorphisms on the risk of colorectal cancer. *Mutagenesis.* 27(2):219-23.
- Toyota M., Ahuja N., Ohe-Toyota M., Herman J.G., Baylin S.B., Issa J.P.** (1999) CpG island methylator phenotype in colorectal cancer. *Proc. Natl Acad. Sci. USA*, 96, 8681–8686.
- Trojan J., Zeuzem S., Randolph A., Hemmerle C., Brieger A., Raedle J., Plotz G., Jiricny J., Marra G.** (2002) Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. *Gastroenterology*;122:211–9.

- Turnbull C.L., Bacon A.L., Dunlop M.G., Farrington S.M.** (2012) In vitro stability of APC gene sequences and the influence of DNA repair status. *Mutagenesis*. 27(2):233-8.
- Vasovcak P., Pavlikova K., Sedlacek Z., Skapa P., Kouda M., Hoch J., Krepelova A.** (2011) Molecular genetic analysis of 103 sporadic colorectal tumours in Czech patients. *PLoS One*. 6(8):e24114
- Vousden, K.H.** (2002) Activation of the p53 tumor suppressor protein. *Biochim. Biophys. Acta Rev. Cancer* 1602:47–59.
- Watson A.J., Collins P.D.** (2011) Colon cancer: a civilization disorder. *Dig Dis*. 29(2):222-8.
- Weisenberger D.J., Siegmund K.D., Campan M., Young J., Long T.I., Faasse M.A., Kang G.H., Widschwendter M., Weener D., Buchanan D., Koh H., Simms L., Barker M., Leggett B., Levine J., Kim M., French A.J., Thibodeau S.N., Jass J., Haile R., Laird P.W.** (2006) CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat. Genet.*, 38, 787–793.
- Welsh K.M., Lu A.L., Clark S., Modrich P.** (1987) Isolation and characterization of the *Escherichia coli* mutH gene product. *J.Biol. Chem.* 262, 15624–15629.
- Wolpin B.M., Mayer R.J.** (2008) Systemic treatment of colorectal cancer. *Gastroenterology*, vol. 134, no. 5, pp. 1296–1310.
- Wood L.D., Parsons D.W., Jones S., Lin J., Sjöblom T., Leary R.J., Shen D., Boca S.M., Barber T., Ptak J., Silliman N., Szabo S., Dezso Z., Ustyanksky V., Nikolskaya T., Nikolsky Y., Karchin R., Wilson P.A., Kaminker J.S., Zhang Z., Croshaw R., Willis J., Dawson D., Shipitsin M., Willson J.K., Sukumar S., Polyak K., Park B.H., Pethiyagoda C.L., Pant P.V., Ballinger D.G., Sparks A.B., Hartigan J., Smith D.R., Suh E., Papadopoulos N., Buckhaults P., Markowitz S.D., Parmigiani G., Kinzler K.W., Velculescu V.E., Vogelstein B.** (2007) The genomic landscapes of human breast and colorectal cancers. *Science*; 318; 1108–1113.
- Zaanan A., Meunier K., Sangar F., Fléjou J.F., Praz F.** (2011) Microsatellite instability in colorectal cancer: from molecular oncogenic mechanisms to clinical implications. *CellOncol(Dodor)*;34(3):155-76.
- Zhang G., Xia S., Tian H., Liu Z., Zhou T.** (2012) Clinical significance of miR-22 expression in patients with colorectal cancer. *Med Oncol*. 2012 Apr 11. [Epub ahead of print]